European Commission

JRC SCIENCE FOR POLICY REPORT

Cross-border transport infrastructure in the EU

A methodology to assess the role of cross-border road networks

Christodoulou A., Christidis P.

2018

Joint Research Centre

EUR 29565 EN

Contact information
Name: Aris Christodoulou
Address: European Commission - Joint Research Centre, Edificio Expo, Calle Inca Garcilaso 3, 41092 Seville, Spain
Email: aris.christodoulou@ec.europa.eu
Tel.: +34 954488276

EU Science Hub
https://ec.europa.eu/jrc

JRC113364

EUR 29565 EN

PDF ISBN 978-92-79-98380-1 ISSN 1831-9424 doi:10.2760/040279

Luxembourg: Publications Office of the European Union, 2018

How to cite this report: Christodoulou A. and Christidis P., *Cross-border transport infrastructure in the EU: A methodology to assess the role of cross-border road networks,* EUR 29565 EN, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-98380-1, doi:10.2760/040279, JRC113364.

Contents

Abstract

The report provides a set of indicators and tools that allow policy makers to measure accessibility and connectivity of border regions in Europe at both national and international levels. The methodology can be used to identify areas where transport infrastructure may be lacking and prioritize potential investments based on specific policy relevant criteria. The approach uses very detailed spatially disaggregate data covering EU28 plus Norway and Switzerland at grid level (1km by 1km), as well as the complete road network. This level of resolution allows many of the specificities of the areas covered to be taken into account.

Acknowledgements

The report has been improved significantly following the valuable recommendations of our Head of Unit, Antonio Soria. It has also benefited from the comments of Carlo Lavalle, Chris Jacobs-Crisioni, Burkhard Schade and Biagio Ciuffo.

The project was completed with the support of our colleagues from the European Commission's Directorate-General for Regional and Urban Policy (DG REGIO), Lewis Dijkstra and Hugo Poelman.

Authors

Aris Christodoulou, Panayotis Christidis (European Commission, Joint Research Centre)

Executive summary

This report provides a set of indicators and tools that allow policy makers to measure accessibility and connectivity of border regions, identify areas where transport infrastructure may be lacking and prioritize potential investments based on specific policy relevant criteria. The approach uses very detailed spatially disaggregate data covering EU28 plus Norway and Switzerland at grid level (1kmx1km), a level of resolution that allows many of the specificities of the areas covered to be taken into account. Data and results can be aggregated at different levels in order to provide a comprehensive view of border zones at different levels of policy analysis.

The European Commission adopted the Communication "Boosting Growth and Cohesion in EU Border Regions" (COM(2017) 534 final) to address the difficulties that border regions face and propose a series of concrete actions to improve cross-border collaboration. The Communication highlights ways in which the EU and its Member States can reduce the complexity, length and costs of cross-border interaction and promote the pooling of services along internal borders. The positive effects of reducing cross-border obstacles can contribute both to the socio-economic development and the integration of border regions. The Commission has a major role in this process by proposing legislation or funding mechanisms or by supporting Member States and regions to better understand the challenges and develop operative arrangements, notably by promoting the information sharing and showcasing successful practices in their border regions.

Transport infrastructure plays a critical role in the connections of border regions and can strongly influence regional, urban or local development. The Trans-European Network for Transport (TEN-T) is of utter importance for the improvement of transport services and cohesion in Europe but it is not sufficient to cover the connectivity deficits of border regions. Border regions are literally in the forefront of geographical cohesion of the EU member states but often appear to be poorly developed and heterogeneous in terms of transport infrastructure. In fact, physical access is the third most cited obstacle (following legal/administrative barriers and language) regarding border regions. The lack of transport infrastructure can be either an issue of natural obstacles, such as rivers or mountains, or an issue of capacity when the existing infrastructure does not cover the high demand because of high interaction levels between the regions sharing a common border.

This report presents a methodology for the analysis of transport infrastructure in border regions assessed in terms of accessibility. The comparison between national and cross-border accessibility helps to capture issues that occur due to the existence of national borders. The main set of results consists of a set of accessibility indicators for each grid cell. The information on grid population and network efficiency can be combined with network data, geomorphology and location of major settlements in order to identify areas in need of transport infrastructure or ways of improving network efficiency. They can also provide information relevant to the economic evaluation of investments. Finally, it is possible to focus on specific grids and examine specific routes in order to identify in detail the parts that can be improved.

Different formulations of accessibility indicators tend to address specific dimensions of the relationship of a point in space with its surroundings. The specific policy question analysed here, i.e. the role of road infrastructure in border regions, can be addressed using three different types of operationalization of accessibility indicators that can quantify three underlying aspects relevant to the spatial relationships of each border region:

- Location indicator as a measure of a border region's connectivity
- Potential indicator as a measure of a region's access to opportunities
- Network efficiency as a measure of the quality of a region's road connections

Aggregating at the level of specific bilateral borders suggests that there is a distinction between older and newer EU member states, central and peripheral European countries or Western and Eastern European countries. At country level, the best performing road network in terms of both national and cross-border network efficiency can be found in the borders of the Netherlands, Belgium and Germany, while the worst performing road network can be found in the borders of the Balkan countries.

The results of the calculation of network efficiency at grid level are used to analyse specific cases. The combination of the results on network efficiency with the road network and population distribution indicates the reasons of low or high network efficiency and highlights the role of the road network. At a more detailed level, network efficiency is combined with data on population of border zone grid cells and destination settlements to identify areas of high interest, i.e. areas with relatively high population and low network performance. The highest concentration of such grid cells can be found in the borders of Czech Republic, but also in the borders between Romania and Bulgaria, Hungary and Slovakia, Hungary and Romania. Using further spatial information it is possible to explore the reasons of low network efficiency. A common reason of low network efficiency is the existence of natural barriers such as mountains (e.g. in the borders between Greece and Bulgaria, Czech Republic and Poland, Italy and France), rivers (e.g. in the borders between France and Germany, Romania and Bulgaria) or national parks (e.g. in the border between Slovakia and Poland). Furthermore, by including the information on both national and international network efficiency the areas in the border between Czech Republic and Poland in highest need for infrastructure investments are identified.

The methodology presented and the resulting dataset can provide useful input to policy makers and practitioners in terms of evaluating the current situation, prioritizing areas for intervention and assessing potential impacts. It should be noted, however, that the areas identified still need to be examined on a case by case basis in order to take into account all the relevant characteristics. The approach allows a focus on specific road links, settlements and physical characteristics, but other factors affecting cross-border collaboration should be also taken into account in additional steps of the analysis.

The overall approach can be extended to cover any area through adapting either the extension of the grid used (e.g. to analyse accessibility of wider zones or even the whole of the EU) or the size of the zone of interest (to calculate indicators using different formulations or attraction factors). The results of this methodology can be easily combined with other policy support tools for tailored multi-criteria analysis.

1 Introduction

The European Union (EU) has 38 internal land border regions in Europe. For historical and geographic reasons, border regions are -in general- more isolated than the rest of the Member State they form part of. On one hand, since most EU Member States have developed their infrastructure and regional policy with a centralised state in mind, the periphery of most Member States has not received as much investment as the core, usually the most central areas of the Member State. In addition, the lack of investment has led to the acceleration of the movements from the periphery to the core, leading to a vicious cycle of further reduced investments in peripheral areas of decreasing importance in population and economic terms. On the other hand, border regions face geographic, historical, cultural and linguistic barriers that limit their opportunities for interactions with their cross-border counterparts which -in most cases- are also isolated within their own national context. These two trends, the internal and the cross-border isolation, mean that a significant part of the EU population has limited access to opportunities, even though they may not be longer considered as "frontier" zones within the European Union.

The European Commission adopted the Communication "Boosting Growth and Cohesion in EU Border Regions" (COM(2017) 534 final) to address the difficulties that border regions face and propose a series of concrete actions to improve cross-border collaboration. The Communication highlights ways in which the EU and its Member States can reduce the complexity, length and costs of cross-border interaction and promote the pooling of services along internal borders. The positive effects of reducing cross-border obstacles can contribute both to the socio-economic development and the integration of border regions. The Commission has a major role in this process by proposing legislation or funding mechanisms or by supporting Member States and regions to better understand the challenges and develop operative arrangements, notably by promoting the information sharing and showcasing successful practices in their border regions. Such actions could enhance the positive integration impacts of EU enlargements as measured in terms of growth of population share in border cities and regions by Brakman et al (2012).

Transport infrastructure plays a critical role in the connections of border regions and can strongly influence cross-border interactions as well as regional, urban or local development. In a study analysing the development of accessibility and population in Western Europe over a 50-year period, Jacobs-Crisioni and Koomen (2017) found that the improvements of the highway network have benefited the improvement of cross-border interaction opportunities more than population has grown. The Trans-European Network for Transport (TEN-T) is of utter importance for the improvement of transport services and cohesion in Europe but it is not sufficient to cover the connectivity deficits of border regions. Border regions are literally in the forefront of geographical cohesion of the EU member states but often appear to be poorly developed and heterogeneous in terms of transport infrastructure. Barriers of major importance imposed by borders include legal ones, as for example relevant to contract enforcement (Rodrik, 2000), as well as linguistic and cultural ones (Persyn and Torfs, 2016). In fact, physical access was the third most cited obstacle (following legal/administrative barriers and language) regarding border regions in a relevant public consultation of the European Commission (European Commission, 2016). The lack of transport infrastructure can be either an issue of natural obstacles, such as rivers or mountains, or an issue of capacity when the existing infrastructure does not to cover the high demand because of high interaction levels between the regions sharing a common border.

The reasons why border regions lag behind in terms of transport infrastructure may have to do with their nature; often they lack the critical mass to justify transport investments (Condeço and Christidis, 2017) or are managed by different authorities which do not always cooperate efficiently. Lopez et al (2009) refer to the issue of 'missing networks' the existence of which – according Maggi et al, 1992 – can be attributed to the fact that for transport development each country tries to address their problem ignoring the

potentials of cooperation and without attempting to coordinate actions. One of the main challenges regarding networks' integration for the European Commission has been to address relevant inefficiencies by contributing to the development of necessary transport infrastructure. As a result, it is important to identify areas in need, or better to classify and benchmark border regions according to accessibility and road infrastructure.

2 Defining accessibility of border regions

The relationship between accessibility and development has been investigated in several cases, especially in the context of assessing major infrastructure investments. There are many definitions and indicators of accessibility addressing different aspects of the topic and covering specific assessment needs. Geurs and van Wee (2004) and Geurs and Ritsema van Eck (2001) provide a thorough review of accessibility indicators based on multiple criteria and taking into account different perspectives. They classify indicators in four main categories: infrastructure-based, location-based, person-based and utility-based indicators. The different categories aim to cover distinct areas of planning or assessment and use information at different levels of detail. The choice of the most suitable indicator depends on a combination of factors including data availability and type of analysis required. Selecting a more informative and detailed indicator is not necessarily the best choice because it might be difficult to obtain the required data while it is important to focus on those factors addressing the issues of interest.

Lopez et al (2008) use four different accessibility indicators to measure how cohesion has changed over the years in Spain. The four indicators correspond to different approaches, focusing on either the location or the infrastructure in combination with different measurement formulations. The population potential indicator or daily accessibility indicator measure reachable population or activities, while the location indicator or network efficiency indicator use the population of destinations as weighting of travel cost measures.

The assessment of network infrastructure and accessibility at NUTS 2[1], NUTS 3 or even municipal level produces relatively rough estimates suitable for analyses at country level or for analysing the impacts of large national or international projects. Gutiérrez et al (2011) used the market potential (an indicator based on the structure of potential accessibility indicators which will be presented in more detail later in this section) to measure the impact of TEN-T projects at NUTS 3 level; they concluded that "the construction of border sections is not very efficient due to the border effect (they produce a lower increase in market potential), though they generate many spill-overs. Internal sections, by contrast, are more efficient, producing more internal benefits but comparatively fewer spill-overs."

There are numerous other formulations of accessibility indicators that can highlight different attributes. In any case, for the identification of the appropriate measure it is important to understand the attributes of the case study area and the priorities of the analysis taking into account data availability. In this study, the main target is to assess the road network of border regions across the EU. By conducting the analysis at fine resolution and for the complete road network it is possible to identify, with relative precision, well performing areas and parts that may require infrastructure investments. Most relevant studies so far have been focusing on specific cases, i.e. countries or regions, and not all Europe while they have been conducted at a relatively coarse spatial aggregation level such as regional or municipal. Furthermore, most studies have been considering accessibility to regional or national capitals. However, for relatively small areas with low population as many border regions are, the attribute of major importance is accessibility to basic services which – in operational terms – corresponds to access to settlements that offer these services or, rather, a selection of settlements/ destinations for each origin-grid to be made by taking into account a measure of attractiveness.

A significant aspect of this study is the coverage of all EU at very fine resolution. Hence, parameters to be considered need to be available at this resolution level. Person-based and utility-based measures are suitable for more local analysis considering characteristics of the activities in the target destinations. Furthermore, neither of the two indicators is suitable for focusing on the road network itself and, most importantly, the

[1] For statistical reasons, data and information is based on NUTS 2 and NUTS 3 regions under the European Nomenclature of territorial units for statistics: http://ec.europa.eu/eurostat/web/nuts

potentials of infrastructure improvement considering geographical, and infrastructure characteristics.

For the analysis of border regions the level of spatial aggregation plays significant role as border regions are often sparsely populated and a fine disaggregation level is necessary to identify areas with inefficient networks and, most importantly, the attributes that reduce networks' efficiency. As the analysis focuses on border regions and the origin points considered are relatively small (1kmx1km grid cells), the potential destinations can be selected among major settlements in relative proximity. Furthermore, working with grids has the advantage of avoiding demarcation problems associated with the selection of nodes or use of nodes as point representations of regions (Bruinsma and Rietveld, 1998). Also, by measuring accessibility at grid level and using the full TomTom network, internal accessibility becomes less of an issue. The network efficiency indicator focuses on the network comparing it to an ideal one by considering travel time and population of one or a number of destinations.

Transport infrastructure of border regions is assessed in terms of accessibility, while the comparison between national and cross-border accessibility helps to capture issues that occur due to the existence of national borders. The main set of results refers to the estimates of accessibility indicators for each grid cell. However, the estimated travel times and distances are also available and can be used for further analysis. The information on grid population and accessibility can be combined with network data, geomorphology and location of major settlements in order to identify areas in need of transport infrastructure or ways of improving accessibility. They can also provide information relevant to the economic evaluation of investments. Furthermore, it is possible to focus on specific grids and examine specific routes in order to specify with detail parts that can be improved.

Different formulations of accessibility indicators tend to address specific dimensions of the relationship of a point in space with its surroundings. The specific policy question analysed here, i.e. the role of road infrastructure in border regions, can be addressed using three different types of operationalization of accessibility indicators that can quantify three underlying aspects relevant to the spatial relationships of each border region:

- Location indicator as a measure of a border region's connectivity
- Potential indicator as a measure of a region's access to opportunities
- Network efficiency as a measure of the quality of a region's road connections

For example, infrastructure-based indicators consider the quality of service (travel time, congestion time etc.) but ignore issues related to activities in the destinations.

Location-based accessibility measures include both a measure of opportunities in the destination zone and a measure of distance or cost of travel between the origin and destination zones. Furthermore, potential accessibility indicators include a cost sensitivity parameter that aims to capture spatial travel behaviour.

Location-based accessibility indicators are the most relevant to the purposes of this study as they can analyse performance of road networks including a measure of importance of destinations.

3 Description of the methodology

The main policy question that the methodology aims to address is how to measure the quality of cross-border road infrastructure. Developing a consistent set of indicators at EU level can assist policy makers in identifying areas for improvement and future investment.

As a first step, the extent of the area to be analysed has to be defined. On one hand, a geographical point is considered as belonging to a border zone if it falls within a distance threshold from an inland border. On the other hand, the area of interest for border zones is considered as a buffer zone, on both sides of the border, where opportunities for interaction may exist. A main point of the analysis is to evaluate and compare the quality of connections and the opportunities that each point in a border zone has with the areas of interest on both sides of the border, the national and the international. The two thresholds – the width of the border zone and the radius of the area of interest – are both user defined. The combination of a border zone distance threshold of 25 km with a radius of 75 km around each point to be analysed encompasses most areas that are of policy interest as regards the potential for cross-border interaction. Both thresholds can nevertheless be modified to reflect different policy focuses. The values selected correspond to zones close to the border for which daily interaction with a cross-border zone is at least theoretically feasible. Changing the thresholds would allow a fine-tuning of the definition of border zones and interest areas, but would also have an impact on the computational complexity and processing time.

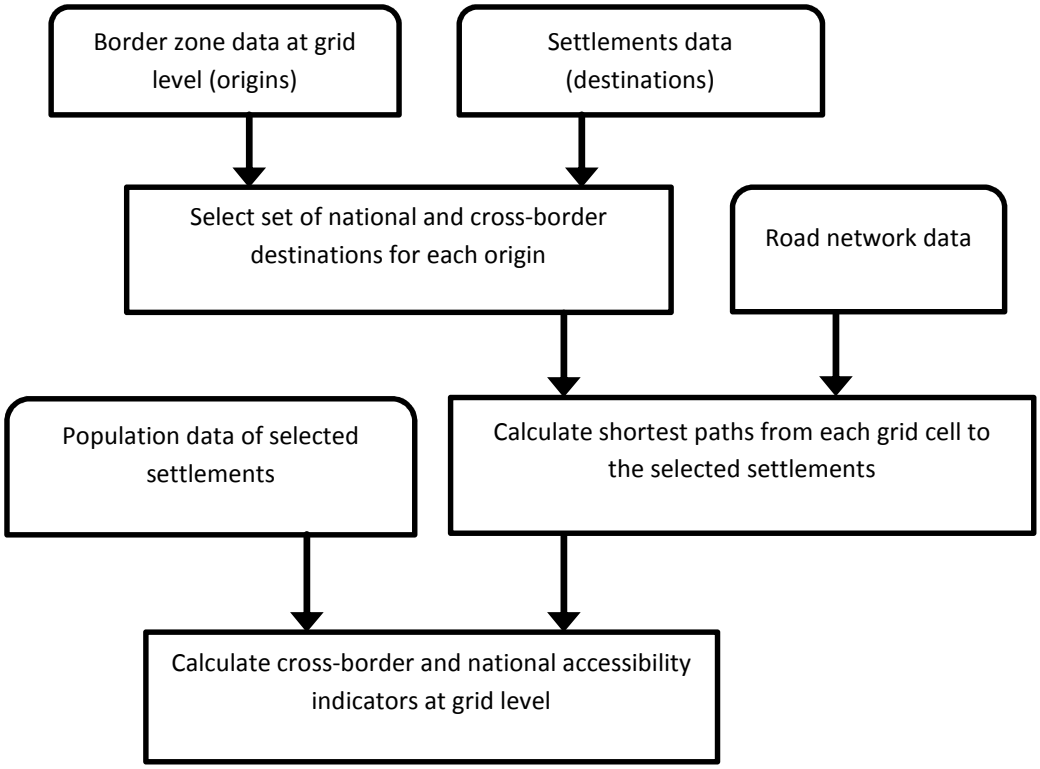
The subsequent steps of the approach combine threes sources of data:

- Population distribution map at 1km by 1km grid (based on the following sources: European Commission (Eurostat, Joint Research Centre and DG Regional Policy - REGIO-GIS) for the areas defined as border zones
- Map of settlements with a population of over 5000 inhabitants (based on the following sources of the European Commission: Eurostat, Joint Research Centre and DG Regional Policy - REGIO-GIS)
- Detailed transport network at European level (TomTom, 2015)

Based on the above, a number of indicators are calculated for each grid cell and the settlements within its area of interest on each side of the border. For the calculation, the following steps have been taken (the calculation process is also illustrated in Figure 1):

- Load location and population data for the 1kmx1km grid cells that lie within a buffer of 25km from each border line between European countries.
- Load location and population data for settlements in Europe with larger than 5000 population.
- Load road network data.
- For each populated grid cell, select the 5 largest settlements within a buffer of 75km in the same country as the grid cell and the 5 largest settlements in the neighbouring country of the grid cell.
- Calculate the shortest path from each grid cell to the selected settlements and keep travel distance, travel time and straight line distance between origin (grid cell) and destination (settlement)
- Calculate cross-border and national accessibility indicators at grid level using travel time and distance information from the previous step, as well as population of the selected settlements.
- Aggregate accessibility indicators using origin (grid cells in border zones) population as weighting.

Figure 1. Flowchart of calculation methodology



4 Exploratory analysis of border zones population characteristics

The first step of the methodology consists of defining the population grid for the interest area. Three parameters were selected in order to match the specification of the policy questions:

- Cell dimension: 1km by 1 km
- Distance from land border: <= 25 km
- Population inside grid: > 0

The third parameter allows the algorithm to discard grid cells that are not relevant for the analysis (since they do not have any inhabitants) and at the same time reduces the number of cells for which calculations need to be made. Most of those non-inhabited cells correspond to areas in the Swedish/ Finish/ Norwegian borders with a particularly low population density.

In terms of population, the interest area is home to almost 13% of the population in the EU, Norway and Switzerland. An interesting aspect of the data is that several grid cells belong to more than one bilateral border area, i.e. they lie within 25 km from the borders with two – and in some cases three – neighbouring countries. The special cases of grid cells belonging to 3 different bilateral borders can be found in Slovenia (401 cells, 2% of the country's total area), Luxembourg (309 cells, 12% of the country's area), Austria (187 cells, including border with Lichtenstein) and France (30 cells bordering Belgium, Luxembourg and Germany at the same time).

The distribution of border regions population among each land border is shown in Figure 2. There are 6.8 million inhabitants within 25 km of either side of the Belgium-Netherlands border and 6.5 million inhabitants along the Germany- Netherlands one. These two borders obviously offer many opportunities for cross-border collaboration, also aided by the cultural and linguistic proximity of the local societies. But it is interesting to note that there are more than 20 other land borders with a population higher than 1 million within 25 km on either side

As a further step for the analysis, the population of border regions can be summarized based on the NUTS 3 zone to which they belong (Figure 3).

Figure 2. Population in border regions, per border

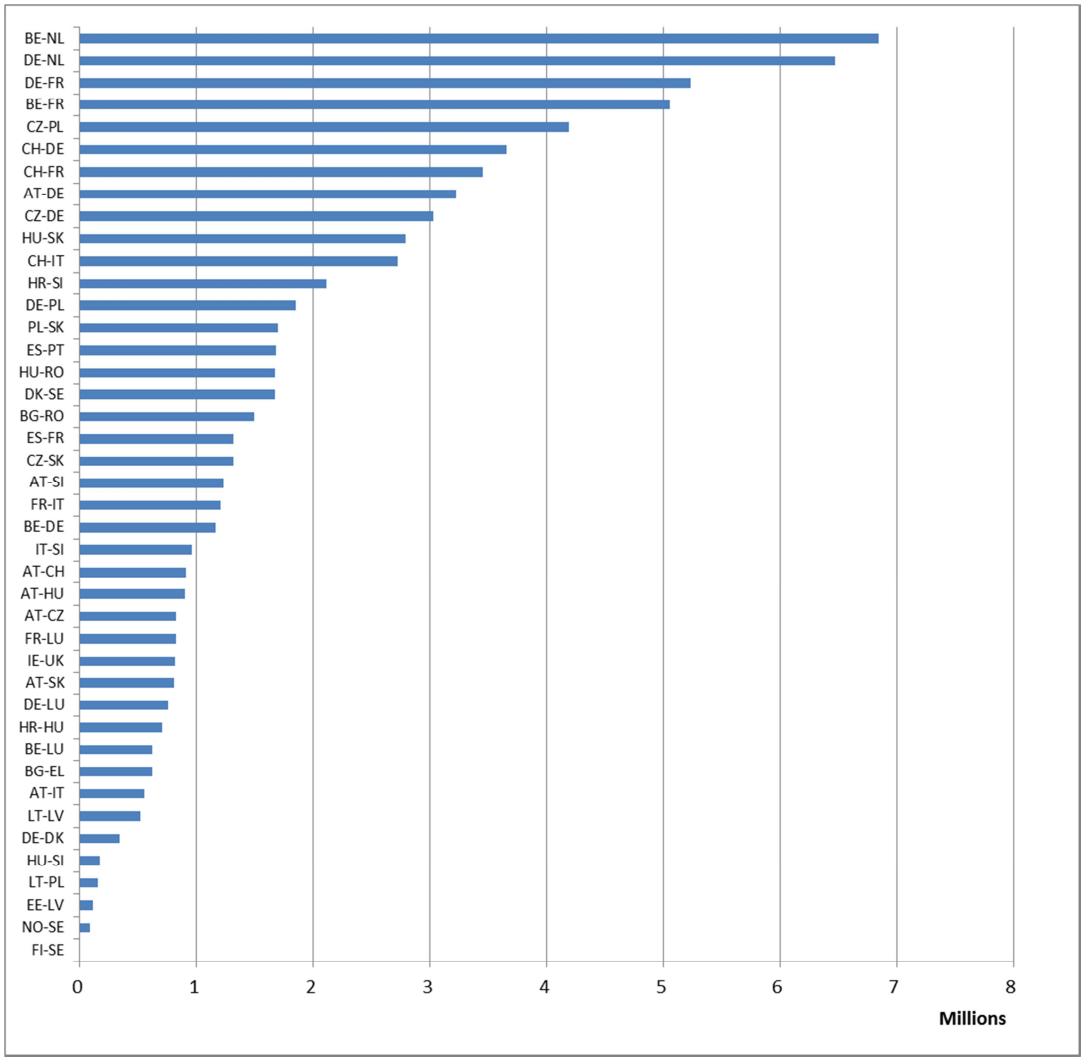
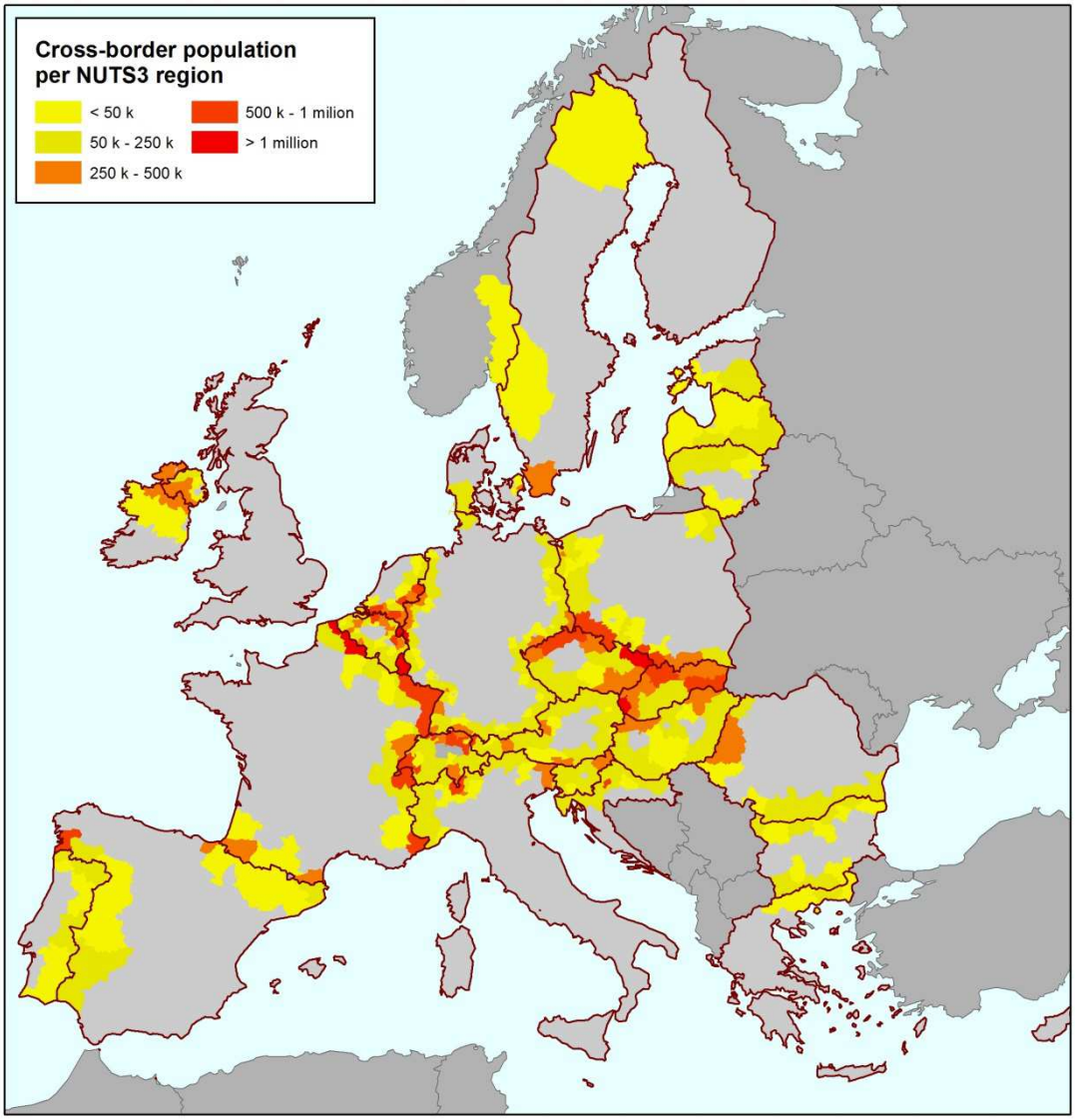


Figure 3. Population in border regions, per NUTS 3 region



5 Measurement of accessibility

As discussed in chapter 2, the analysis of accessibility of border areas requires a combination of indicators in order for the various policy aspects to be taken into account. The following indicators have been calculated:

- Location indicator
- Potential accessibility indicator
- Network efficiency indicator

Each indicator on its own allows a comparison between the national and international accessibility of each zone, but the pairwise combinations of the three indicators permit further insights into the reasons of the observed situation and –in particular- the evaluation of the potential impacts of improvements in the road network on cross-border accessibility.

5.1 Location Indicator

The first indicator calculated is the location indicator, which measures travel time weighted by the population of destinations. It is given by the following formula (Gutiérrez et al, 1996):

$$LI_i = \frac{\sum_{j=1}^{n} t_{ij} P_j}{\sum_{j=1}^{n} P_j}$$

This indicator should be interpreted from a locational perspective as it represents the travel time between a location and a number of points of interest (Gutiérrez, 2001), where:

t_{ij} the travel time from cell i to settlement j

P_j the population of settlement j

n the number of points of interest (destination zones or settlements j) to be taken into account in the calculation

Travel time is calculated over the road network, using the route assignment algorithm described in chapter 6.

Population P has been selected as the attraction/weighting factor for the operationalisation of the indicator since it is readily available and is a good proxy for most economic and social interactions. Nevertheless, several other attraction/ weighting factors can be chosen for the calculation of alternative location indicators if data are available (economic data, number of businesses, hospitals, etc.) to complement this population-based location indicator.

The number of settlements to take into account was set to 5, i.e. the location indicator refers to the weighted travel time to the five largest settlements over 5 thousand inhabitants within a threshold distance of 75 km from the grid cell. Both the number of 5 settlements and the 75 km threshold were selected empirically as a sufficient approximation of the policy objectives that define this study. Using fewer than 5 settlements may lead to unstable results that over-rely on specific settlements, while using more increases computation time without any meaningful change in the results (a test carried out with 10 indicators suggests that the values do not change significantly).

The indicator is expressed in time units, and its physical interpretation is how long on average it takes to drive from each cell to the main activities within 75 km. Travel times have been calculated separately for destinations within the same Member State as the cell and for destinations in the neighbouring Member State, still within a 75 km radius from the grid cell.

The median of the location indicator at national level is 49 minutes (i.e. for half of the cells more than 49 minutes are needed to reach the main activities). The median on the international side rises to 59 minutes. The median of the ratio between the international and national side is 1.2, i.e. 20% longer travel time on the international side than on the national side (Figure 4). It is worthwhile to note that a large number of cells have a ratio lower than 1, meaning that the travel time to activities in the neighbouring Member State is lower than in the country where they belong to. On the other hand, while at national level most activities can be reached within 1.5 hours, the distribution at international level has a *longer tail*, with a larger share of cells over the 1.5 hours mark.

Figure 4. Distribution of location indicator values per cell (national, international and ratio of international to national)

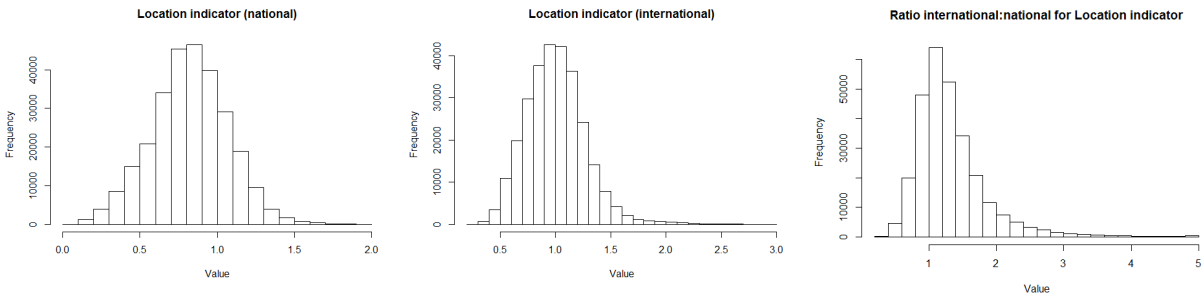
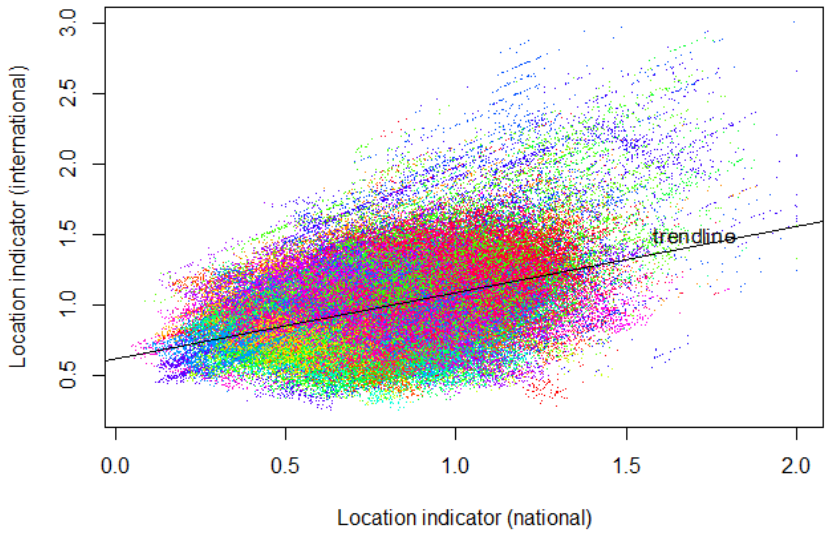


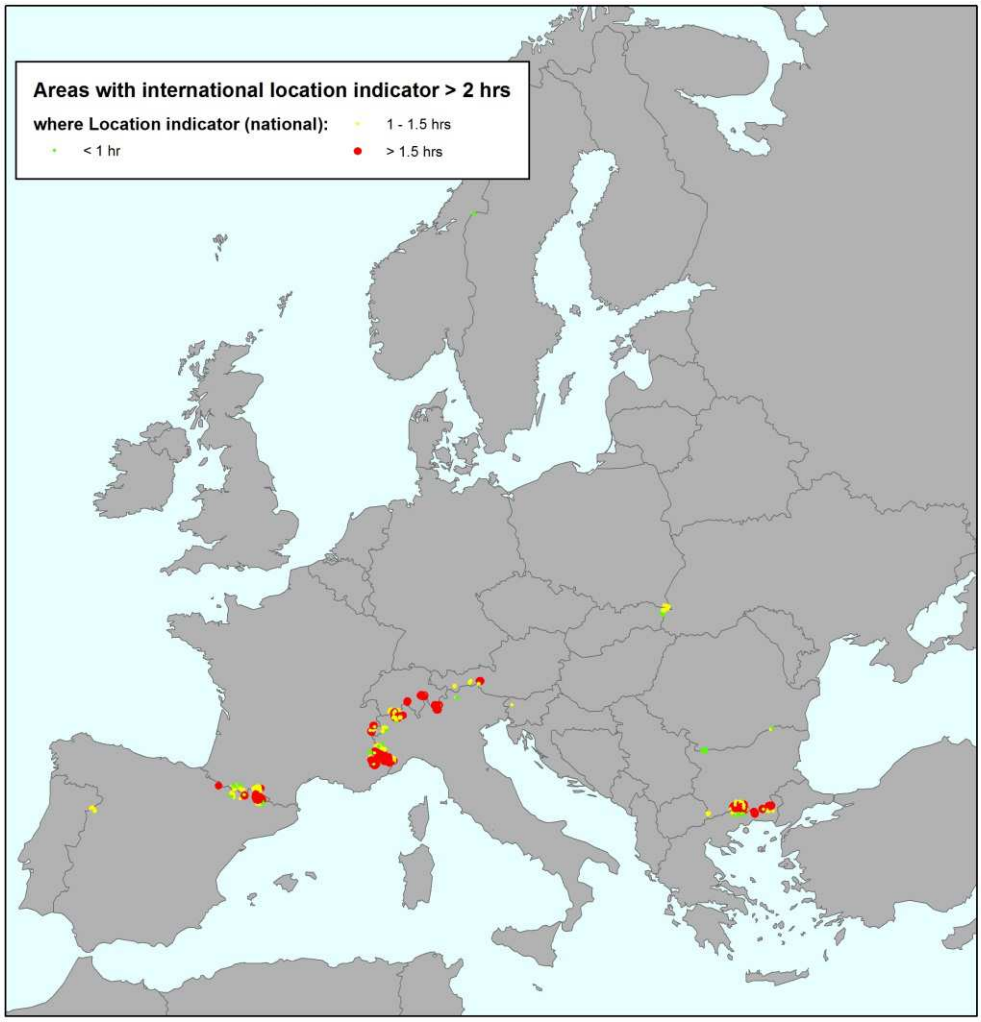
Figure 5 shows the correlation of the international and national location indicators per cell. The graph implies a roughly linear relation, with the intercept at 0.6 (which in physical terms would correspond to an average distance to the border of 36 minutes). The different colours in the graph correspond to different bilateral borders (as described in chapter 4). It is evident that there is a certain clustering effect that can be attributed to specific geographic conditions, such as natural barriers or lacking network links.

Figure 5. Correlation of national and international location indicator



As an example, the zones with an international location indicator higher than 2 hours can be seen in Figure 6, with a comparison of the national location indicator for the same cells. Cells with high values on both sides of the border (in red) are probably isolated because of the morphology of the area, such as the Alps, the Pyrenees and the Rodopi ranges. The points on the Bulgarian-Romanian border are the result of the lack of river crossings over the Danube.

Figure 6. Areas with international location indicator higher than 2 hours (colours correspond to the national location indicator for the same cell)



5.2 Potential Accessibility indicator

The second indicator is potential accessibility, which is classified as a gravity-based indicator considering the way it incorporates travel time. Travel time is raised to the power of *a* in order to control the importance of the role of distance or time between origin and destination. Values larger than 1 (i.e. a higher decay function parameter) increase the importance of relations over short distances (Gutiérrez, 2001). For the results presented here the power parameter used is 1.

$$PA_i = \frac{\sum_{j=1}^{n} P_j}{\sum_{j=1}^{n} {t_{ij}}^a}$$

where:

t_{ij} the travel time from cell i to destination zone (settlement) j

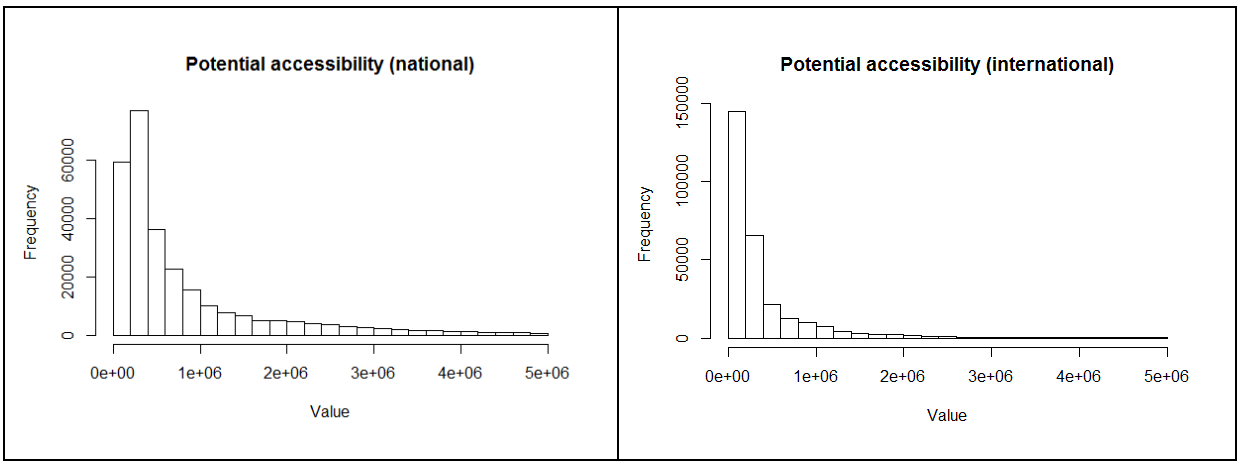
P_j the population of destination zone j

n the number of points of interest (destination zones or settlements j) to be taken into account in the calculation

a is a parameter to control the decay function

This indicator should be interpreted from an economic perspective as it measures the economic potential of each place considered and the changes to be caused by new infrastructure (Gutiérrez, 2001). As in the case of the other indicators, population P is used as a proxy of the opportunities that can be accessible. The simplest formulation uses parameter a=1, which corresponds to a linear decay function. The value of the resulting indicator for each cell is proportional to the sum of population accessible from the cell per unit of travel time (e.g. hour). Since a constant radius of 75 km around each cell is used, the cells farthest from the border obviously have higher accessibility to population in their national territory rather than to population in the neighbouring one. As a result of this geometric property, the distribution of the Potential Accessibility Indicator differs significantly between the national and international level (Figure 7).

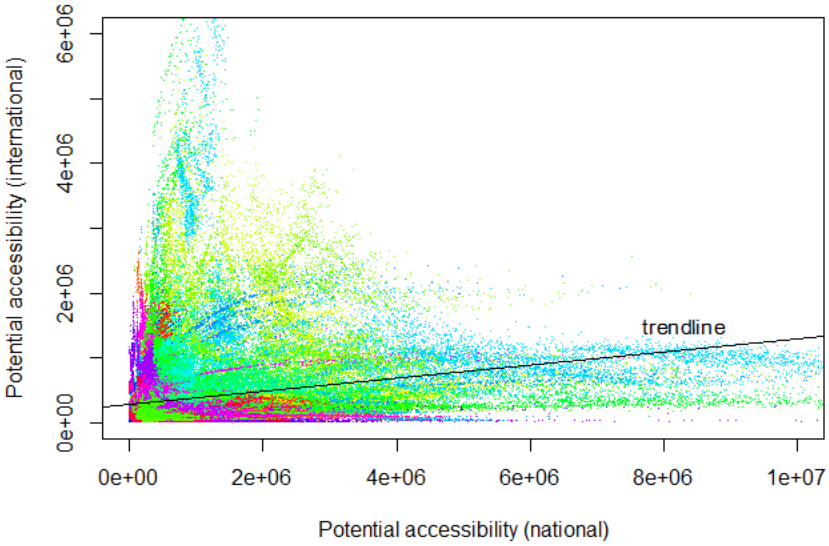
Figure 7. Distribution of Potential Accessibility indicator values per cell (national and international level)



For the national part, the mean Potential Indicator is 923 (thousand persons per hour), with the median being 416 k per hour (the long tail of the distribution causes high skewness). For the international part the figures are much lower, with a mean of 379 k per hour and a median of 191 k per hour. The correlation between the value of the indicator at national and international level can be seen in Figure 8, using different colour for cells belonging to the same border. A mixed picture is produced, with a large number of cells having low values for both. Those cells probably correspond to rather isolated cells close to the border. Two other distinct groups can be identified: the

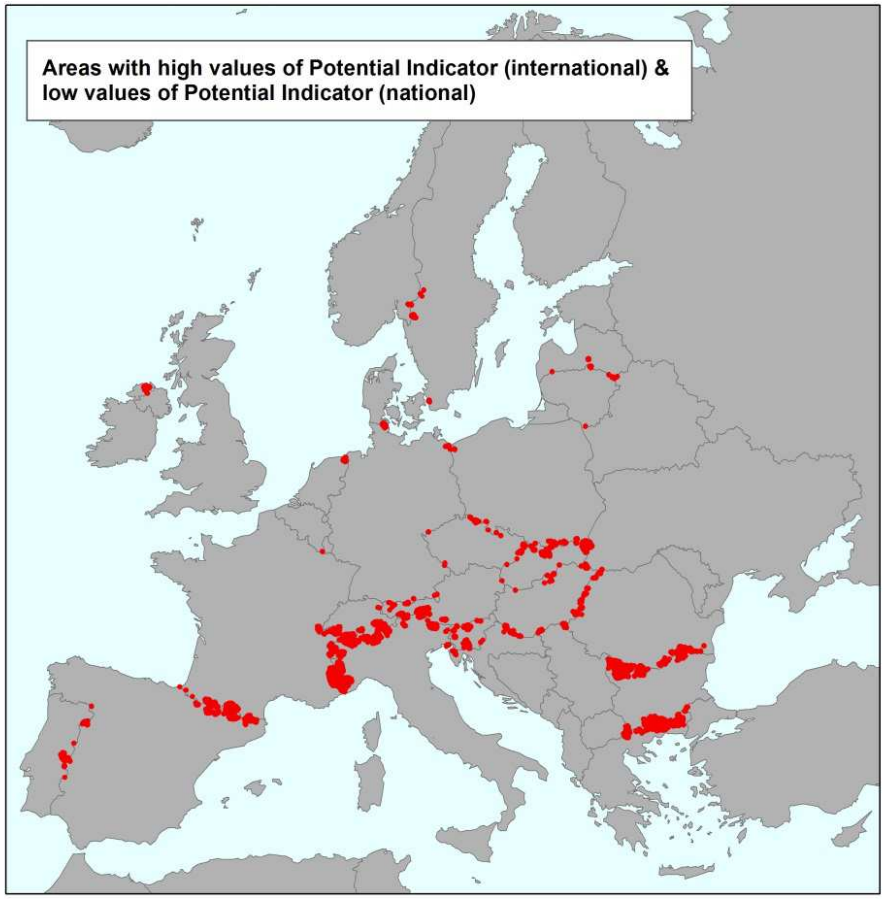
majority corresponds to cells with high values for national and low values for international (zones farther from the border with no significant levels of population in the neighbouring country), but an interesting group from the policy perspective consists of the zones with low values for national and high values for international Potential Accessibility.

Figure 8. Correlation of national and international level Potential Accessibility indicator per cell



Filtering the results with minimum and maximum values for the national and international side of each cell allows a focus on specific typologies of border areas. For example, Figure 9 maps the zones with high international and low national Potential Accessibility. The thresholds used in this example were a minimum of 2 million on the international side, maximum of 2 million on the national side.

Figure 9. Areas where Potential Accessibility is high at international level and low at national level



5.3 Network Efficiency

Network efficiency indicators were initially proposed by Gutiérrez and Monzon (1998) as an elaboration of the 'route factor' used to measure the sinuosity of individual links. The rationale is to reduce reliance on geographic location – or to "neutralise the impact of the geographic location" – and to assess transport infrastructure needs or potentials of a region. The network efficiency indicator considers accessibility on the existing network and accessibility on an ideal network in order to indicate the potential of improvement of the network.

The network efficiency indicator NER, which is reported in the results of this study, is the reverse of Network Efficiency (NE) which is given by the following formula:

$$NE_i = \frac{\sum_{j=1}^{n} \frac{t_{ij}}{t'_{ij}} P_j}{\sum_{j=1}^{n} P_j}$$

I.e. NER is given by the following formula:

$$NER_i = \frac{1}{NE_i}$$

where,

t_{ij} is the travel time on the network from origin zone i to destination zone j

t'_{ij} is the time needed to cover the straight line distance between origin zone i and destination zone j when travelling with a constant speed of 120km/h

P_j is the population of destination zone *j*

n is the total number of destination zones to be considered, in the specific case n=5

As t_{ij} is the ideal travel time, NE_i will tend towards 1 as accessibility improves and will increase as accessibility reduces. By reversing network efficiency the NER_i indicator obtains values between 0 and 1, so that values reaching 1 indicate high network efficiency while values reaching 0 low.

In Gutiérrez and Monzon (1998) instead of weighting the ratio of travel times by population, it is weighted by the incomes of the economic centres in the destinations. Economic activity might be more representative of the attractiveness of a settlement as a destination than population. However, it is very difficult to obtain reliable data of economic activity, especially at the scale required, and population is directly related to the attractiveness of a destination.

In general, the scale of application – all Europe analysed at grid level, in the specific case – dictates the choice of indicators, variables and methods. The combination of population data at grid level, data regarding settlements and a very detailed road network makes this analysis unique. At the same time processing data of such detail and size is quite challenging. In any case, the assessment of network and connectivity of border regions using the network efficiency indicator weighted by population, has produced some very interesting outputs.

The network efficiency indicator takes into account attributes of the network by including a variable representing ideal travel time. The ideal travel time incorporates issues regarding both road length and speed as it refers to travel time for a straight line distance between two points under a maximum speed. As a result, the speed has to do with the type of the road and the length with the design that may be influenced by several factors including the geomorphology.

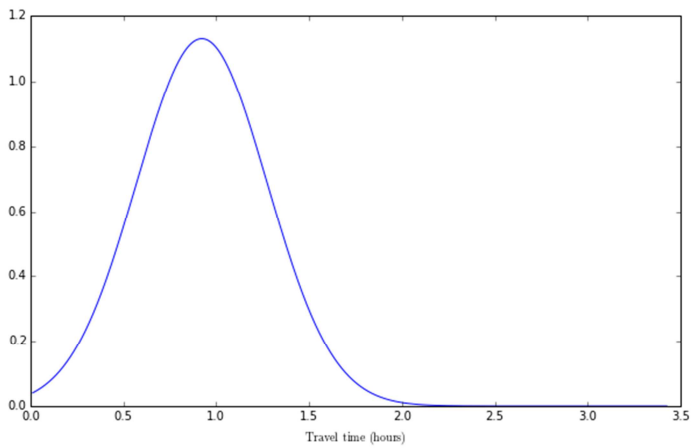
The network efficiency indicator has been identified as the most suitable one to assess and benchmark road infrastructure of border regions. In addition to performance of the network (travel time) and attractiveness of potential destinations (population) it includes a parameter representing ideal performance of the network. As a result, the indicator focuses on the characteristics of the network, remains unaffected by the physical distance between origin and destination and uses population to weight travel time. Before elaborating on the results of network efficiency, the indicator is compared with the two other accessibility indicators presented earlier in order to highlight their main differences and showcase its ability to better capture the characteristics of road infrastructure.

National and international network efficiency is calculated separately for each couple of countries sharing borders. More specifically, for each origin zone (grid cell) national network efficiency is calculated considering the 5 most populated settlements in the same country within a radius of 75km and international network efficiency is calculated considering the 5 most populated settlements in the neighbouring country within a radius of 75km.

The impact of limiting the total number of destinations to 10 (5 national and 5 cross-border) has been examined by considering accessibility to 20 in total destinations (10 national and 10 cross-border). The disadvantages of increasing the number of destinations have to do mainly with the processing time which practically doubles when doubling the number of shortest paths to be calculated. Furthermore, the consideration of 10 (5 national and 5 cross-border) destinations from each grid cell is absolutely sufficient to capture the attributes of the network. The comparison of the run with 10 in total destinations with the run using 20 destinations showed small differences in the values of network efficiency: in only 4% of the cases for which network efficiency is calculated the difference is larger than 10%.

Regarding the consideration of destinations, a radius of 75km has been used to take into account settlements within relative proximity. Working at fine resolution level makes it possible to capture with higher precision the potential destinations of interest. As a result, reality is represented better in comparison to considering only major cities. The means of the travel distances and travel times to the selected settlements are 68km and slightly below 1 hour (55 minutes), respectively. The distribution of all the travel times calculated is presented in Figure 10.

Figure 10. Frequency distribution of estimated shortest path travel times



The points of departure are the centroids of the inhabited 1kmx1km grid cells within a 25km buffer of national terrestrial borders and the points of arrival are the centroids of settlements with population larger than 5000 inhabitants. More specifically, travel times between the nodes of the road network that are closest to the origin and destination

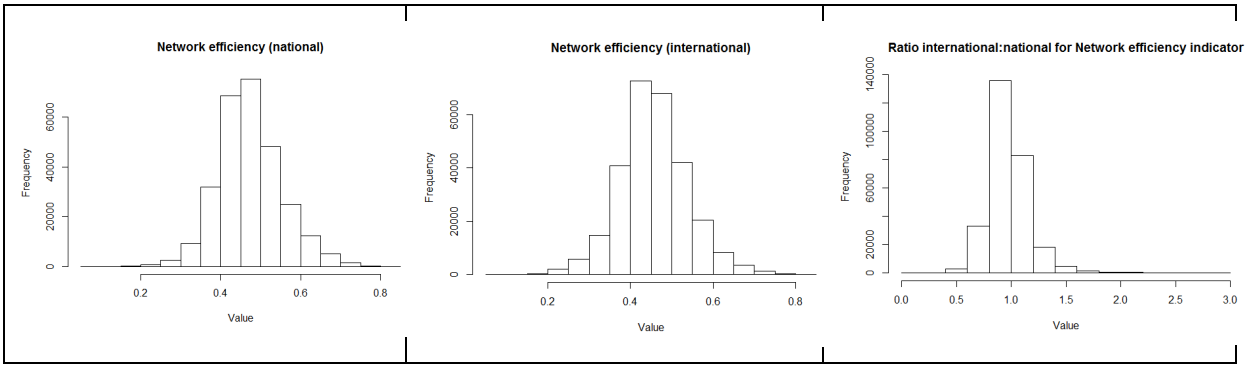
points are calculated. Travel times are based on the calculation of the shortest path using the A* algorithm and the speed information included in the TomTom data for each link that is part of the shortest path. The full TomTom road network is used.

Only populated grid cells are considered, while in a few cases of sparsely populated areas no settlements could be found within the determined buffer. These cases are not considered when reporting the total results of network efficiency. A map of grid cells without settlements – either in the country where the grid cell is located or in the neighbouring one – within the determined radius (75 km) is presented in Figure 11. The majority of cases of grid cells without settlements in the buffer can be found in the Swedish-Finish, Norwegian-Finish and Norwegian-Swedish borders, all very sparsely populated areas. Other areas with high concentration of grid cells without settlements within the determined buffer are in the borders between Switzerland and Italy, Italy and Austria, Austria and Czech Republic, France and Spain, Spain and Portugal, Estonia and Latvia, all in sparsely populated areas.

Figure 11. Grid cells without settlements in a 75km buffer

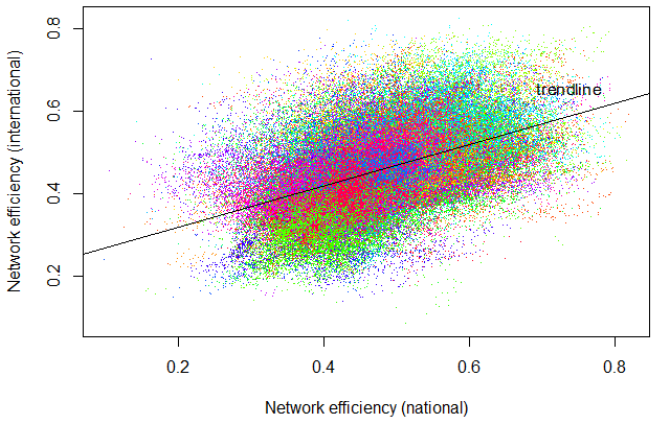
The values of the network efficiency indicator vary significantly depending on the local conditions. Figure 12 shows the distribution of the indicator for the national and international sides of the borders, as well as the distribution of the ratio of the two indicators. The median values of the national and international side are relatively similar, 0.47 and 0.45 respectively, but the variance for the international part is higher. The higher values for international network efficiency correspond to areas very close to border crossings served by highways, while the lowest values to zones that are located close to natural obstacles. In the first case, the morphology and the road design permit a travel time close to the theoretical travel time at maximum speed at straight line. But in the second case multiple factors affect the low value of the indicator: large distance to border crossing, inefficient road design due to difficult morphology, low operational speed.

Figure 12. Distribution of network efficiency (national and international) and ratio international to national



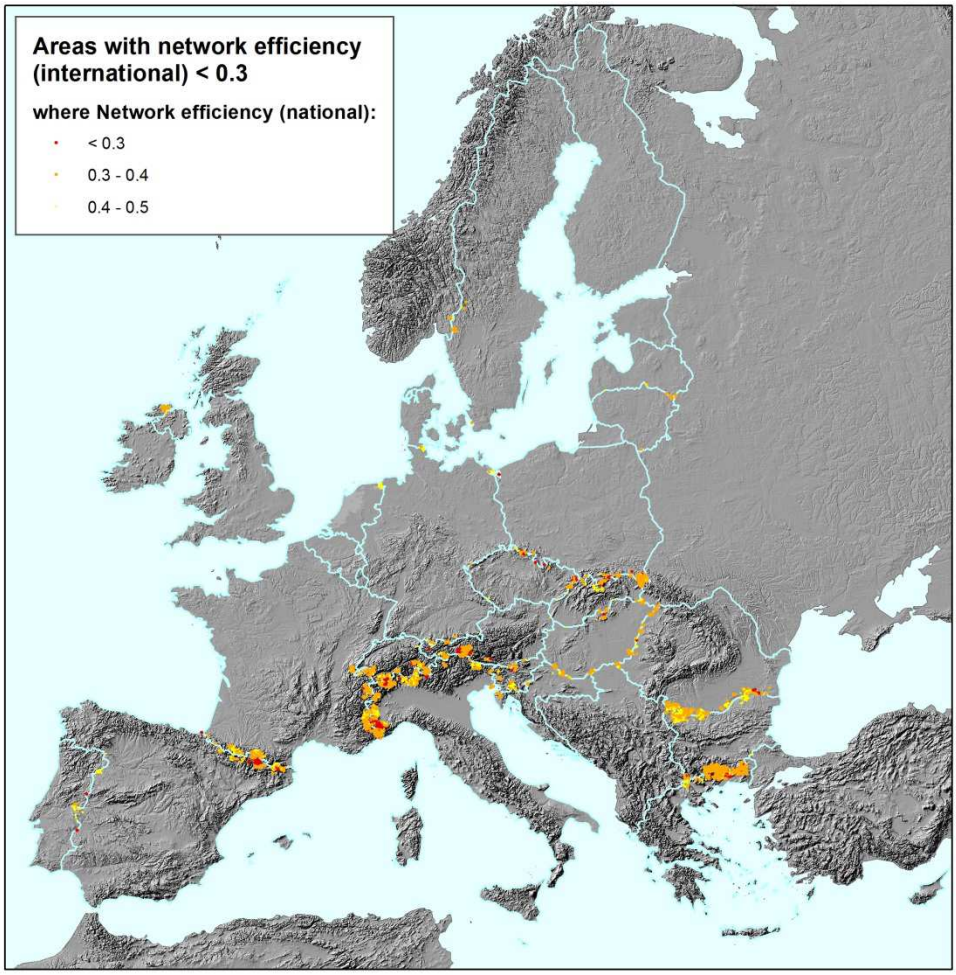
The correlation between the national and international side (Figure 13) suggests that there is a rough linear relation between the two. This implies that, in general, the geographic conditions that affect network efficiency have a similar impact on both sides of the border. This would mean that where the two values are similar (ratio ≈ 1) physical obstacles are probably the reasons for a low network efficiency indicator. Using different colours for each border (Figure 13) also reveals the existence of several clusters in the data, further reinforcing the hypothesis that a large part of the value of the network efficiency indicators depends on local conditions.

Figure 13. Correlation of national and international level Network Efficiency indicator per cell



The concentration of zones with low values for network efficiency on both sides of the border is evident in Figure 14. Virtually all cells with a low value (<0.3) on the international side have a value lower than 0.5 on the national side as well. The combination of the location of those zones with the elevation map of Europe confirms the hypothesis of a high concentration on specific borders. Moreover, most of the zones with values lower than 0.3 on both sides of the border (red point on map) are located in mountainous areas.

Figure 14. Correlation of national and international level Network Efficiency indicator per cell (source of elevation map: European Environment Agency[2])



[2] https://www.eea.europa.eu/data-and-maps/data/digital-elevation-model-of-europe

5.4 Comparison of the three accessibility indicators

The results of the calculation of the three indicators (location indicator, potential accessibility and network efficiency) for national and cross-border accessibility, aggregated at border region level, are presented in Figure 15. The results of network efficiency can be found in the left hand side column, those of the location indicator in the middle and the potential accessibility calculations are presented in the right hand side column. In the upper row are presented the results regarding accessibility within the country and in the lower row are presented results referring to cross-border accessibility.

The three indicators are presented together for comparison reasons. The aggregation at the level of border regions is suitable for this purpose as main differences between them are shown. In order to compare the three indicators some further processing of the calculations is required because they represent different aspects of accessibility and their output is not directly comparable.

The location indicator measures travel time weighted by the population of the destination. In this case population represents the attractiveness of a zone and is used to weight travel time, while accessibility is measured in terms of travel time, i.e. the smaller the location indicator the higher the accessibility.

On the other hand, the potential accessibility indicator measures the number of accessible population considering travel time. Adopting a gravity approach, the population is used as a measure of attractiveness the intensity of which decreases with distance, travel cost or travel time. Hence, in this case the larger the potential accessibility indicator the higher the accessibility.

Finally, the network efficiency indicator has a structure similar to the location indicator but takes into account also ideal travel time. Hence, accessibility is relative to ideal travel time and improves as travel time approaches the ideal one.

The indicators in Figure 15 are scaled from 0 to 1 and the results are classified in four equal intervals (0-0.25, 0.25-0.5, 0.5-0.75, 0.75-1) in order to illustrate the best (darker) and worst (lighter) performing regions according to each indicator.

Potential accessibility is heavily influenced by distance or travel time, but in this case the relevant impact is expected to be relatively limited as a result of limiting the search of destinations to the settlements within a radius of 75km from the origins. The results of potential accessibility are to a large extent driven by the size of population of destinations. This has been confirmed by obtaining a very similar pattern to the one presented in the right column of Figure 15 from the aggregation of the destination population by border regions. This explains the unexpectedly high results in some regions, such as on the border between Romania and Bulgaria, but also indicates a weak point of the indicator. The general picture is affected by the fact that by aggregating the results at the level of border regions the most populated grid cells dominate the totals. Nevertheless, the facts that the indicator is significantly influenced by the population of the destinations and does not focus on the role of the road network make it unsuitable for the type of analysis required.

A projection of the three indicators in 3 dimensions allows more messages, as regards their comparisons, to be derived (Figure 15 to Figure 19). The three indicators present low levels of correlation, both at national and cross-border level. The cross-border side of all three indicators has a higher variation than the respective indicator at national level. Nevertheless, while high levels of network efficiency are uniformly spread across various levels of potential accessibility and location indicator levels at national level, the cross-border distribution is much more unbalanced. A relatively small number of cells presents high network efficiency, high potential accessibility and low location indicator (for which low values mean low average travel times) on either side of the border, with the majority corresponding to the national side. These zones correspond to border regions with easy access to highways, are close to large settlements on the other side of the border and have access to a considerable size of population within 75 km. This is the

case mainly for zones on the Netherlands-Germany border. On the other extreme, a relatively large number of zones is considered isolated according to any of the three indicators (i.e. lower values of network efficiency and potential accessibility, higher values of location indicator). The share of zones with such values at national level is considerably lower, confirming the hypothesis that in general cross-border accessibility is lower than accessibility at national level.

Figure 18 and Figure 19 add a fourth dimension, using a different colour for each land border. It is worth noting that clusters of zones belonging to the same border can be easily observed, probably due to the similarity of the conditions across the whole border, that affect (at least) potential accessibility in a similar way. Potential accessibility appears to be the variable that allows the clustering per border to be visible both at national and cross-border level. It also appears that the ordering of the border clusters is roughly the same for both sides of the border, the difference being that the distance between the clusters is much lower at national level. The physical interpretation of this observation is that all three indicators are affected by common factors related to a specific land border, which to a large extent has a similar effect on zones on either side of the border. For example, zones on both sides of the border in a mountainous area are expected to have low values of network efficiency, while border zones located in low density areas would need long travel times (high location indicator) to access major settlements on either side of the border.

Figure 15. Accessibility indicators. 1st col: network efficiency, 2nd col: location ind., 3rd col: Potential acc./ 1st row: national, 2nd row: cross-border

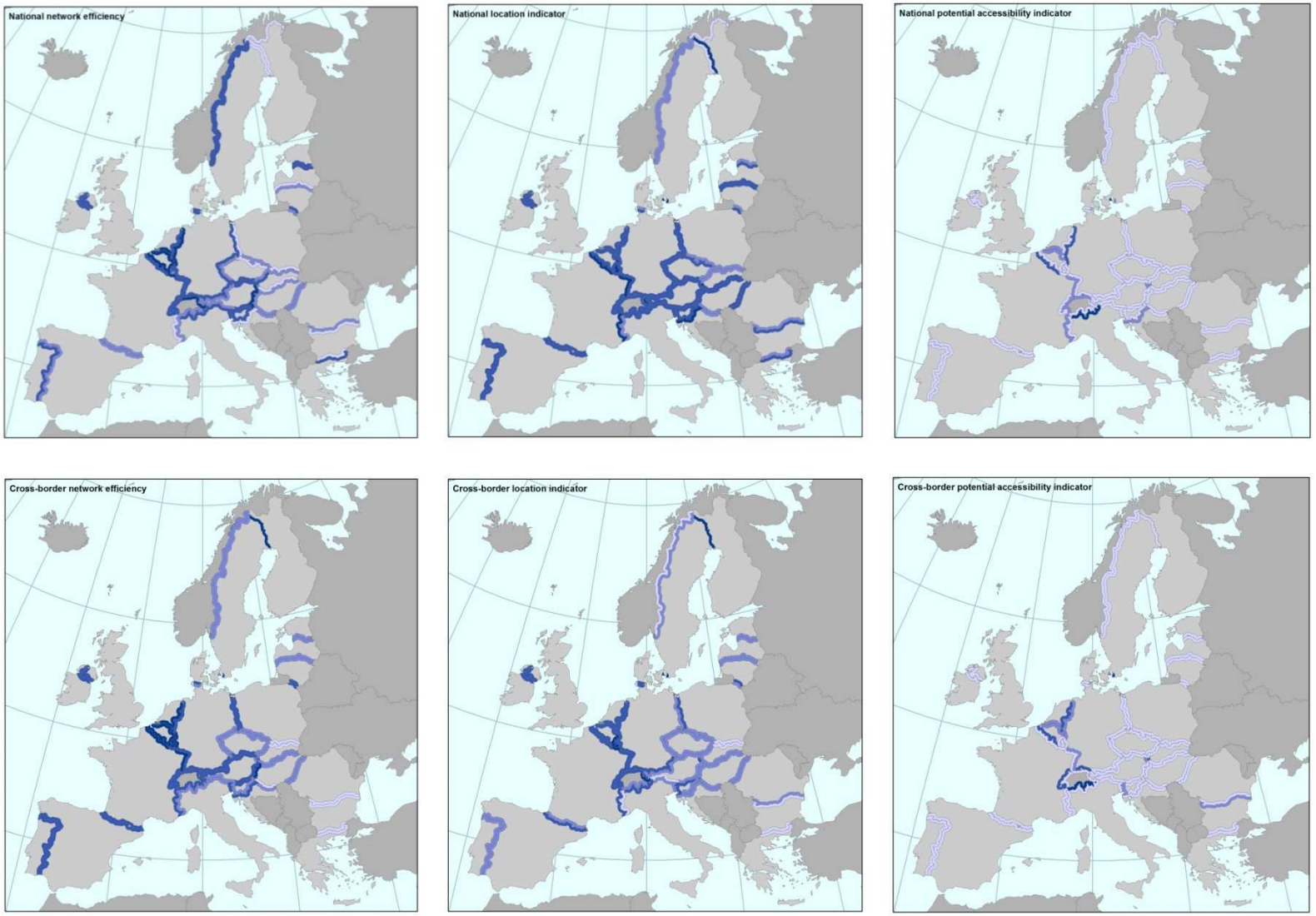


Figure 16. Correlation graph of the three indicators, cross-border

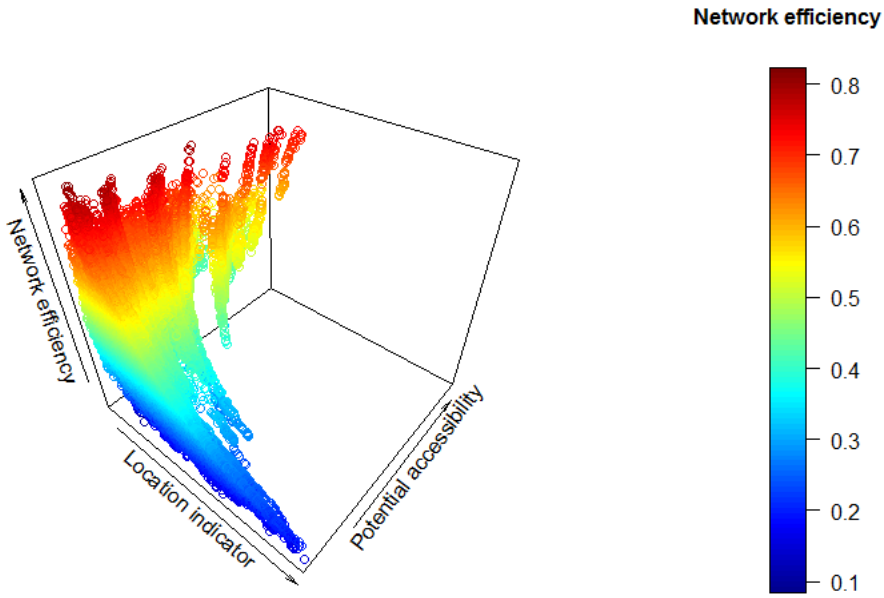


Figure 17. Correlation graph of the three indicators, national

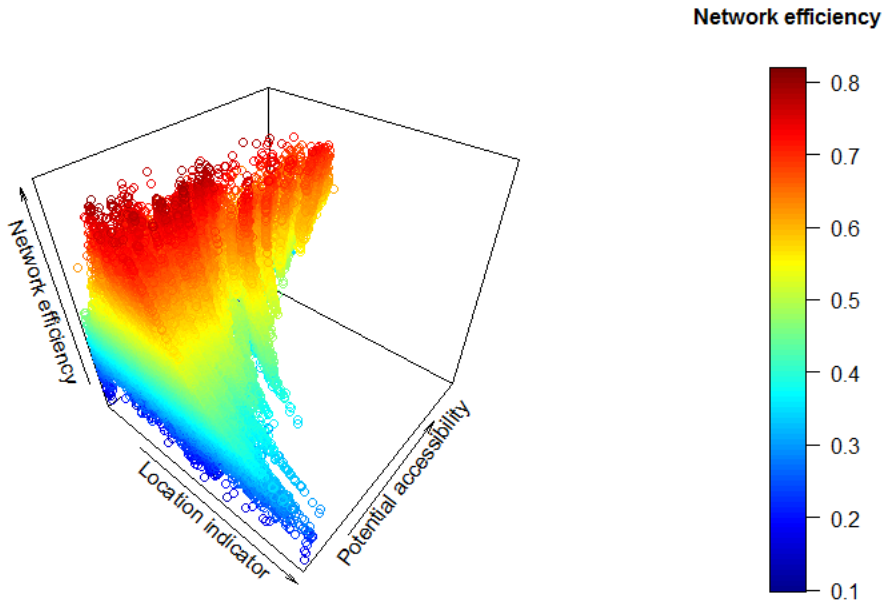


Figure 18. Correlation graph of the three indicators, cross-border, by bilateral border

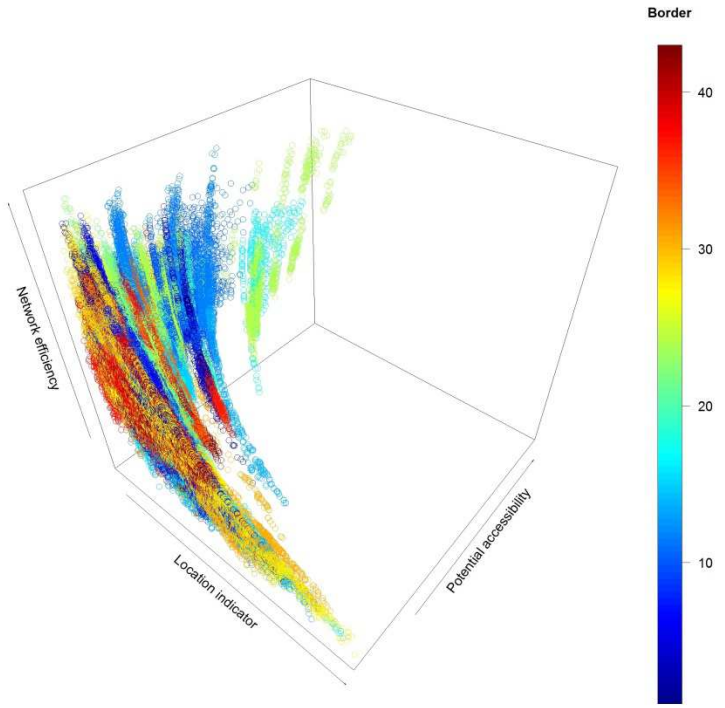
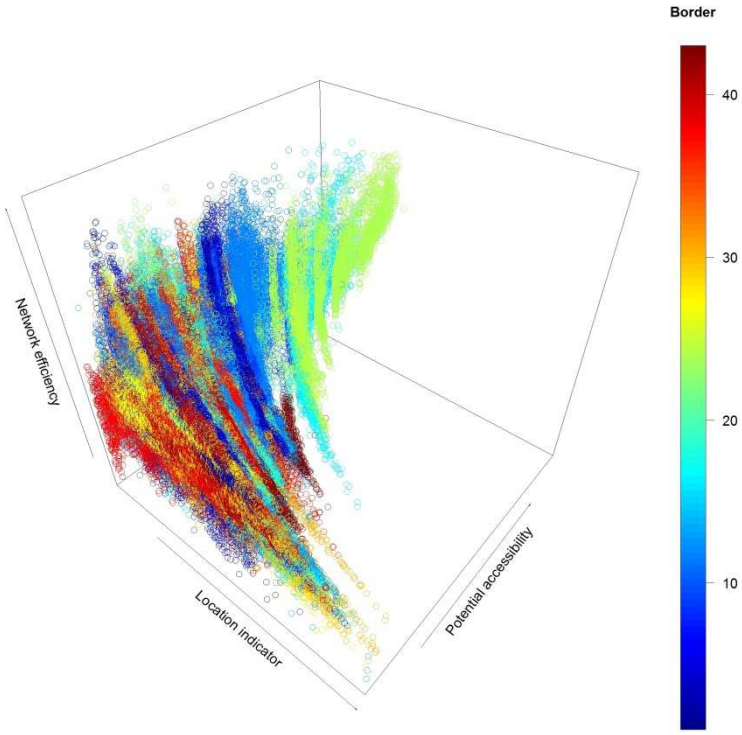


Figure 19. Correlation graph of the three indicators, national, by bilateral border



6 The shortest path algorithm

The most computationally intensive process of the approach is the calculation of shortest paths between the grid cells of border zones and the settlements within a radius of 75 km. Path search algorithms, starting from a point, expand the search to neighbouring nodes in the graph and select the nodes to be added to the path that comply with certain criteria, most commonly cost minimisation. Calculation time increases as the area (part of the graph) covered during the search increases. A heuristic can be used to limit the number of nodes to be considered in order to avoid wasting time with nodes for which it is obviously impossible to be part of the optimal path. On the other hand, if the heuristic makes the algorithm ignore nodes of the shortest path then the calculated path will not be the optimal one.

The Dijkstra algorithm guarantees the calculation of the shortest path between two points. Starting from the origin, the Dijkstra algorithm gradually covers all alternative routes until reaching the destination; this is when the search stops. The shortest path is selected as the one that minimises the cost function g(n). Without directing the path search towards the destination, the Dijkstra algorithm covers routes in all directions and as a result, especially in big networks, it can take long time to reach the destination. This is also the case of the present application where a very detailed network is used and many shortest paths are calculated.

The A* algorithm is capable of calculating routes relatively fast. Similarly to the Dijkstra algorithm, the A* algorithm bases the selection of routes on a cost function g(n) (from origin to node n) that might be referring to travel time, distance etc. The main difference with the Dijkstra algorithm is that the A* algorithm includes also a heuristic that directs the path search towards the destination in order to reduce the part of the network covered and as a result the calculation time. The heuristic function refers to a cost estimate h(n) from any node n to the destination, which is minimised when closest to the destination. During the path search this cost estimate h(n) is added to the cost function f(n) in order to direct the search towards the destination. Finally, the A* algorithm finds the path that minimises $f(n)=g(n) + h(n)$.

The heuristic function can affect not only the processing time but also the final output and for this reason it must be determined very carefully. For the determination of the heuristic, a compromise is made between running speed and precision. The heuristic should be of the same scale as the cost function. A heuristic higher than the cost function from the node to the destination will improve processing time but does not guarantee calculating the shortest path. A heuristic lower than the cost function will always lead to the shortest path but the lower it is, the slower the algorithm will become as the search will cover bigger part of the network. Furthermore, the heuristic function must be as simple as possible considering that it will be calculated as many times as the number of nodes considered.

In this study, a heuristic that does not always lead to the optimal path has been determined in order to achieve a processing time that permits the completion of the exercise. However, the divergence of the results from those using a heuristic to calculate optimal paths is very small and definitely disproportionately small in comparison to the impact that the use of such a heuristic would have on processing time. From a small sample examined, the divergence of the accessibility indicator as calculated using the determined heuristic in comparison to the results obtained using a heuristic suitable to calculate optimal paths is larger than 10% for 4% (or larger than 20% for 0.2%) of the border region grid cells considered. Furthermore, the impacts of this divergence would be further reduced considering the variation of the limits used to present the results of the accessibility indicator.

For the calculation of network efficiency, and most importantly of the shortest paths the programming language python and network processing libraries have been used.

In order to limit the memory requirements, the road network is loaded partially so that to cover the wider border area of the couple of countries considered. The processing time for a specific couple of countries depends on the number of shortest paths calculated but also on the length of these shortest paths, or better the number of links included. For the 43 couples of countries considered, more than 2.5million shortest paths have been calculated when 5 national and 5 cross-border destinations are considered.

7 Aggregate analysis of border regions

The primary outputs from the calculation of network efficiency are (a) the shortest paths table including the travel times and distances between each origin (grids) and destination (settlements) and (b) the network efficiency table including the NER indicator for each grid. These results are further processed, aggregated and mapped to produce meaningful and easy to interpret outcomes.

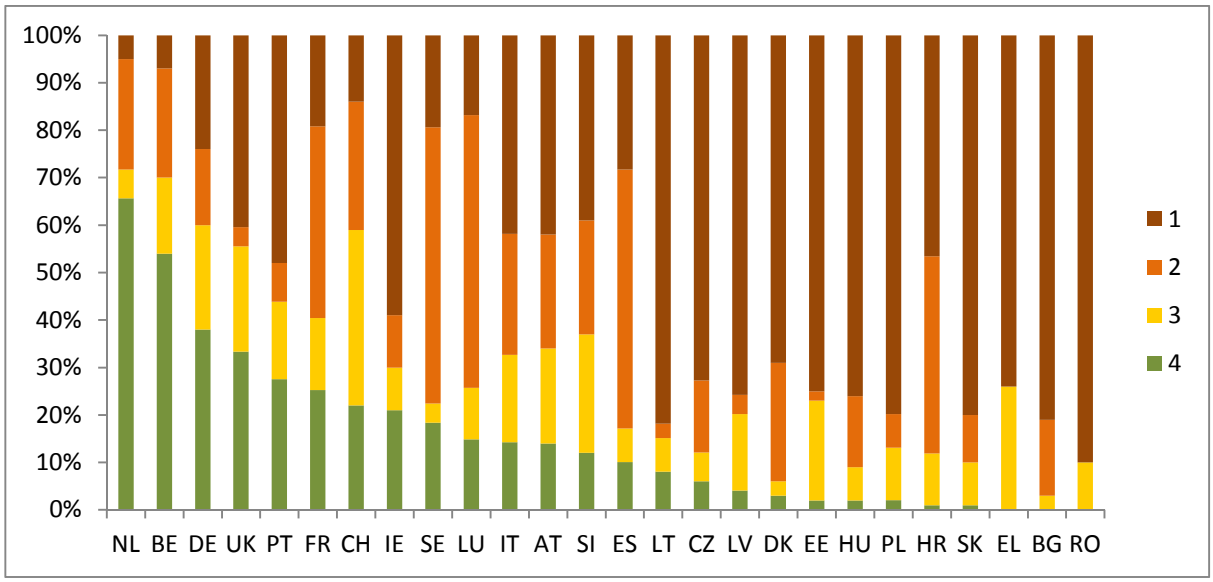
Average network efficiency of border regions is a first indicator of the quality of the network. Although the identification of specific areas in need or the determination of necessary measures do require thorough examination of the characteristics of the area, aggregate results provide a general view of the situation in the EU. Furthermore, the difference between national and international network efficiency can help to make a first short listing of the areas in potential need of investments in infrastructures.

At first national and cross-border network efficiency are combined to rank countries in terms of the proportion of the population affected by the quality of the network in border zones. As network efficiency (NER) takes values between 0 and 1, each grid cell can belong to one of the following four categories:

1. National network efficiency < 0.5 and cross-border network efficiency < 0.5
2. National network efficiency < 0.5 and cross-border network efficiency ≥ 0.5
3. National network efficiency ≥ 0.5 and cross-border network efficiency < 0.5
4. National network efficiency ≥ 0.5 and cross-border network efficiency ≥ 0.5

Obviously, category 1 refers to the worst performing grid cells, while grid cells in category 4 are performing well in terms of both cross-border and national network efficiency. Having classified border grid cells in one of the four categories, the population of each grid is used to find the distribution of the border regions' population according the network efficiency level of their residential location. The results are aggregated at country level and Figure 20 presents a ranking of the countries in terms of the network efficiency to which is exposed the population living in border regions. For countries with more than one neighbour, cross-border efficiency to all neighbours is considered.

Figure 20. Distribution of border population according to the level of network efficiency



The Netherlands is the country with the highest proportion of border population (more than 60%) with access to well performing national and cross-border network followed by Belgium (more than 50%) and Germany (almost 40%).

Interesting outcomes immediately observed in the maps of national and international network efficiency (first column of Figure 15) refer to differences between older and newer EU member states, central and peripheral European countries or Western and Eastern European countries. The most efficient EU networks considering both national and cross-border connections can be found in the borders between France, Belgium, the Netherlands, Germany and Luxemburg. On the other hand, the worst performing cross-border EU connections are found in the East of the EU and more specifically in the borders between Greece, Bulgaria, Romania, Poland, Slovakia.

Other issues to be observed in the national and cross-border network efficiency maps refer to specific cases. For example, among the Baltic countries Latvia seems to have the most inefficient road network. Considering the LT-LV border zone, the national network efficiency of Latvia is in the lowest band. This is the only border zone in the area of the Baltic countries with low performance as the border zones in the EE-LV and LT-PL have relatively high network efficiency.

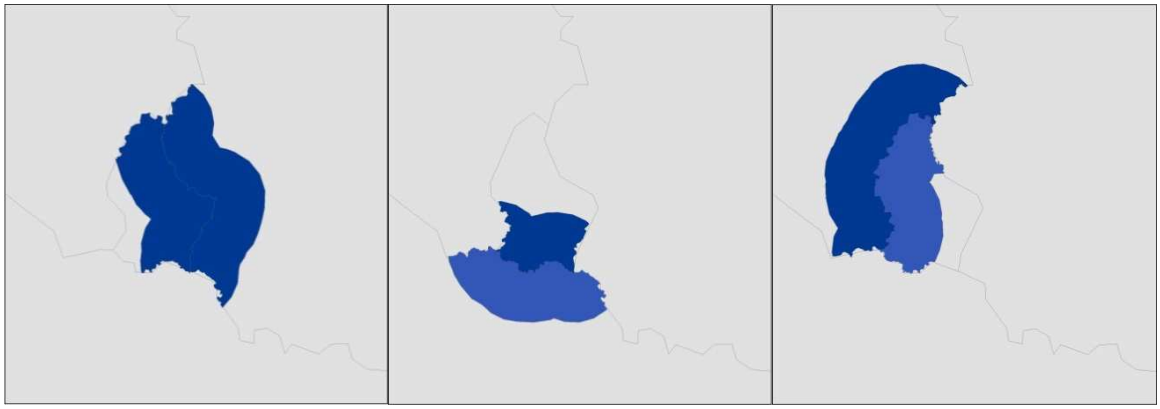
The network efficiencies of both border zones of Bulgaria (with Greece and Romania) are in the lowest band and this might be explained by the existence of natural borders: mountains between Bulgaria and Greece and the Danube river in the borders with Romania.

Another case to be examined in more detail is the network of the Iberian Peninsula which has high cross-border network efficiency in both border zones of Spain with France and Portugal.

The border zone between Poland and Slovakia will be examined in more detail as cross-border network efficiency appears to be in the lowest band for both countries.

In the maps presented in Figure 15 several border zones overlap. Although the general picture is presented, each couple of countries can be examined separately. In Figure 21 a closer view to the case of Luxemburg is presented. More specifically, cross-border network efficiency of Luxemburg and its three neighbouring countries (Germany, France and Belgium) is presented. According to the results, cross-border NER of Luxemburg to and from Germany is relatively high, cross-border network efficiency from Luxemburg to France is more than 0.6 and cross-border NER from France to Luxemburg below 0.5.

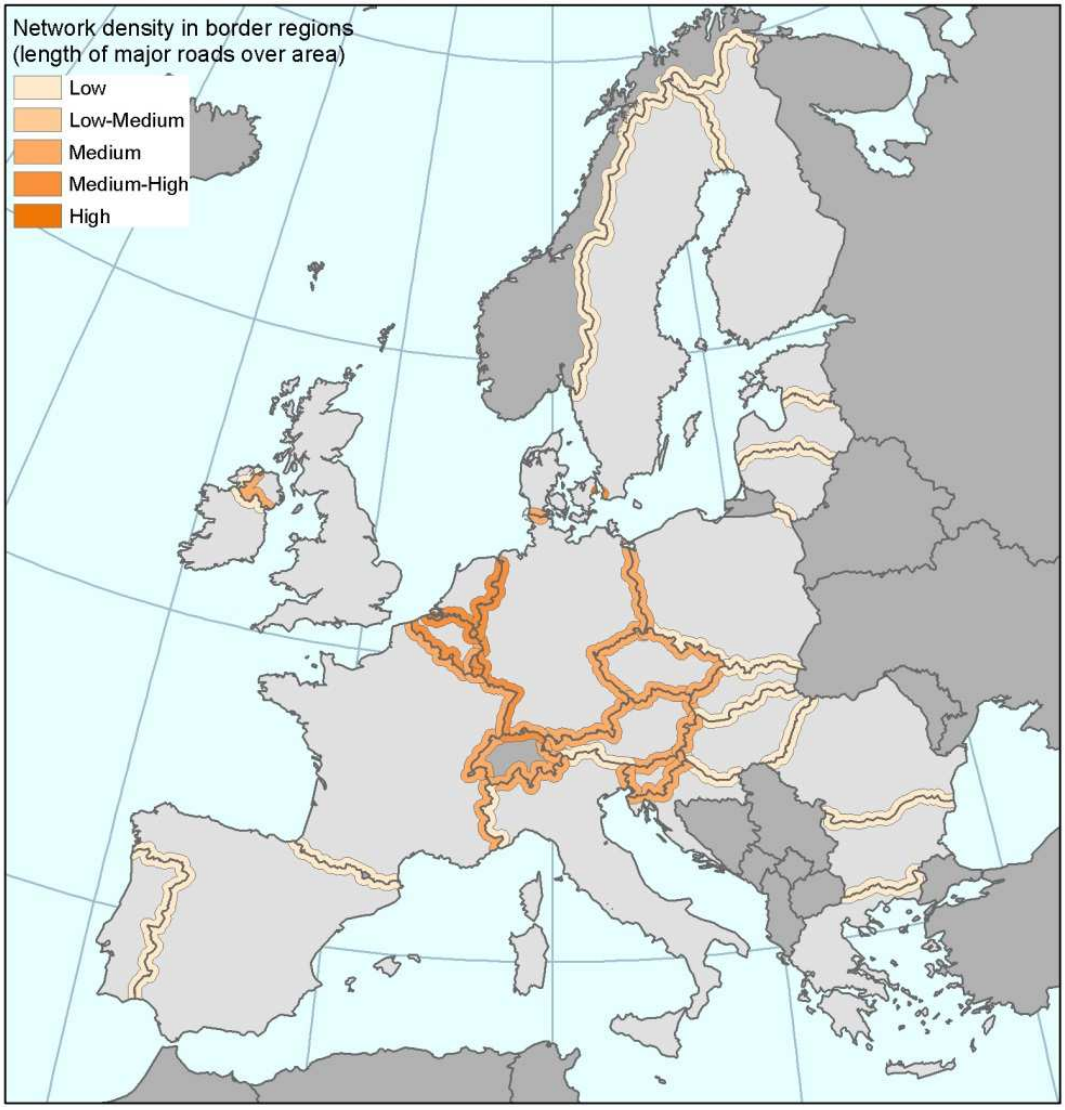
Figure 21. Cross-border network efficiency of Luxemburg and neighbouring countries (on the left: Germany, in the middle: France, on the right: Belgium)



By examining closer the case of Belgium it can be noticed that cross-border network efficiency between Belgium and its neighbouring countries is higher than national. The relatively high difference between Germany and Belgium can probably be attributed to the relatively less efficient network in Belgium in comparison to Germany. The reason is that although cross-border network efficiency for the couple Germany-Belgium is more or less the same for both Belgium and Germany (around 0.54), for Belgium national network efficiency is slightly lower (0.53) than cross-border, while for Germany it is higher (0.57).

Network efficiency is a measure of how well the network operates and network density is often a determining factor of performance. In Figure 22 the density of major roads in border regions is presented. The estimates presented have been produced using the first three Functional Road Classes (FRCs) of the TomTom data and they refer to length of roads (meters) by border area (squared meters). This map presents very clearly the distinction between central and peripheral countries. However, when combined with the network efficiency maps of Figure 15 some interesting observations can be made. For example, in the borders between Spain and Portugal and Spain and France network density is relatively low but network efficiency relatively high. This might indicate that the network serves effectively the needs in these areas. Similar observations can be made for the border zones between Latvia and Estonia, Lithuania and Poland.

Figure 22. Density of major roads in border regions



8 Grid level analysis

The main strength of the approach followed in this study is that the analysis is conducted at a highly disaggregated spatial level with respect to the full road network. The results at grid level in combination with other attributes of the area can indicate regions that suffer from poor accessibility and play a significant role in recommending measures. The maps in Figure 23 and Figure 24 present the results of national and cross-border network efficiency, respectively, at grid level for Europe.

Poorly performing areas as shown in these maps include the borders between Greece and Bulgaria, central part), Bulgaria and Romania (western part), France and Spain. However, each border area should be analysed separately in order to determine infrastructure needs or identify the main attributes that affect network efficiency. Several examples will be examined in the rest of the report. Most of them have been selected based on the aggregate results of border regions.

Figure 23. National network efficiency at grid level

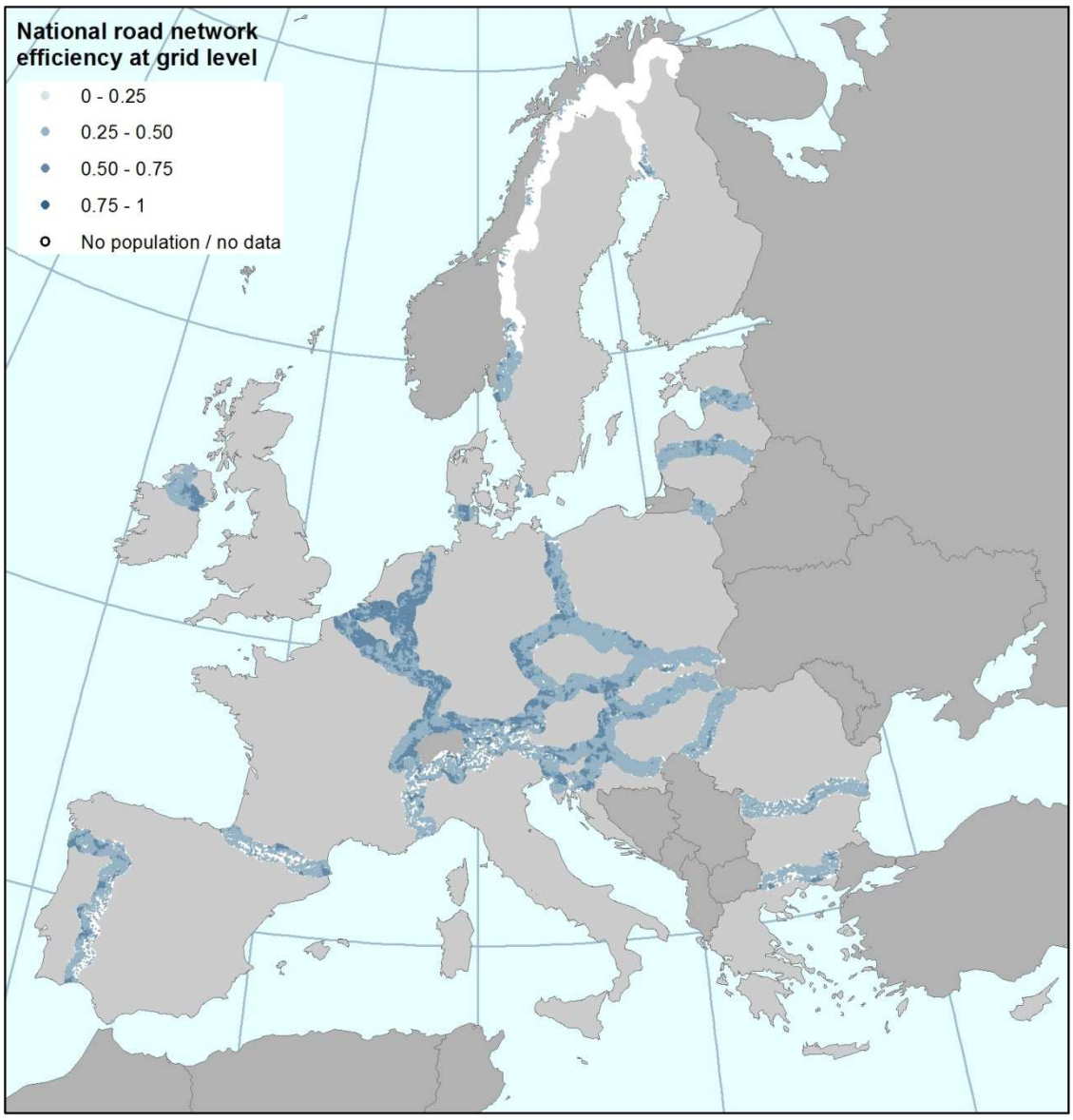
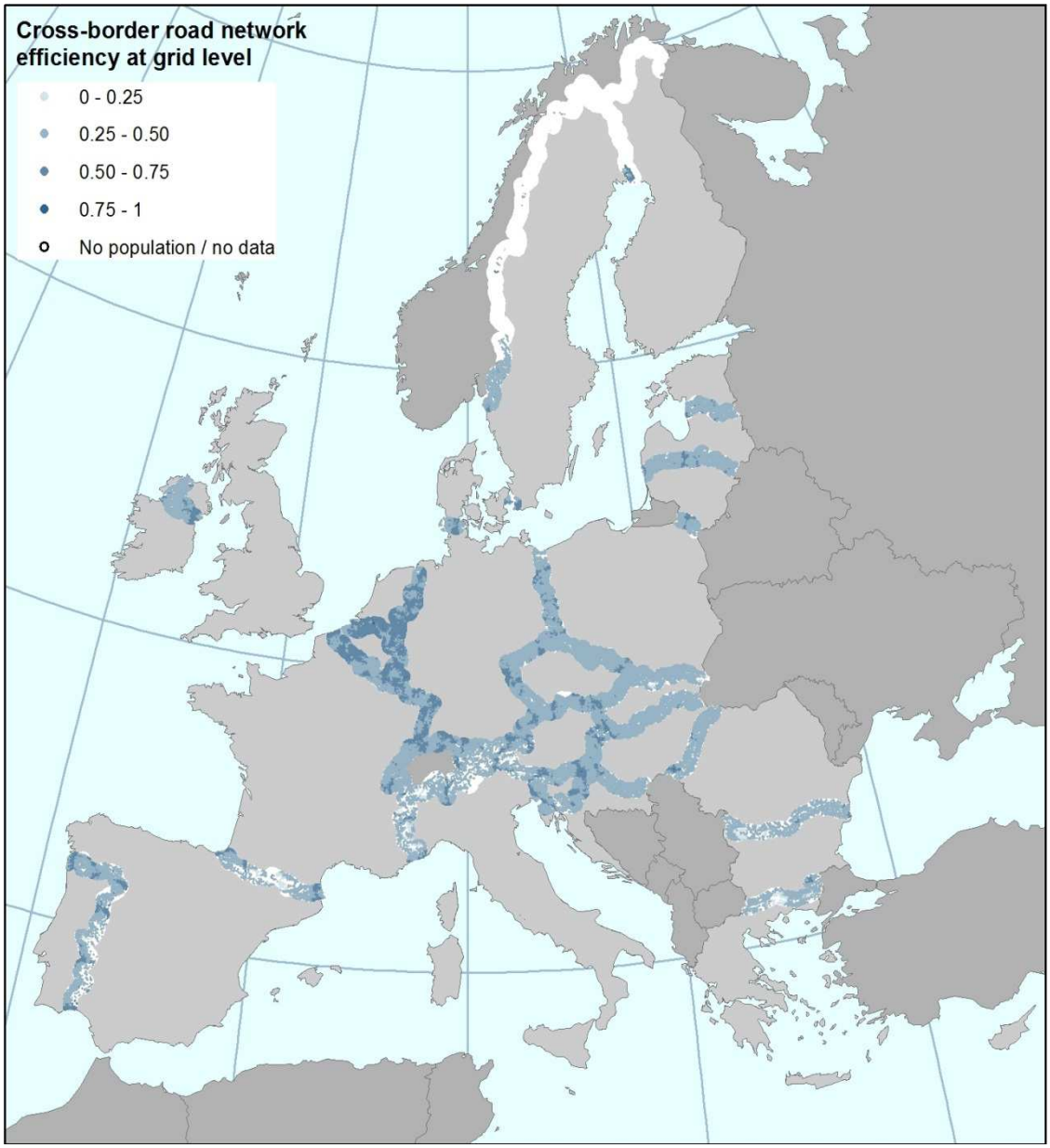


Figure 24. Cross-border network efficiency at grid level



8.1 Border regions in the Baltic area

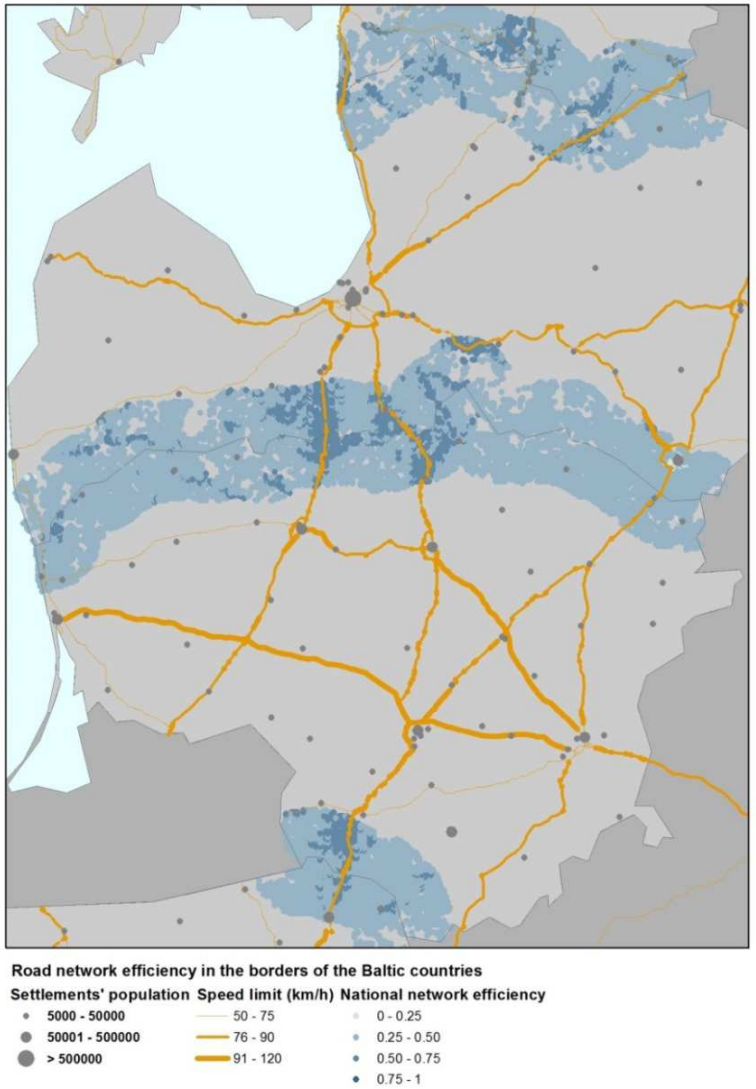
The results of national and cross-border network efficiency at grid level for the Baltic countries are presented in Figure 25 together with the settlements of more than 5000 people, as well as the network of motorways and major roads according to the classification of the TomTom data.

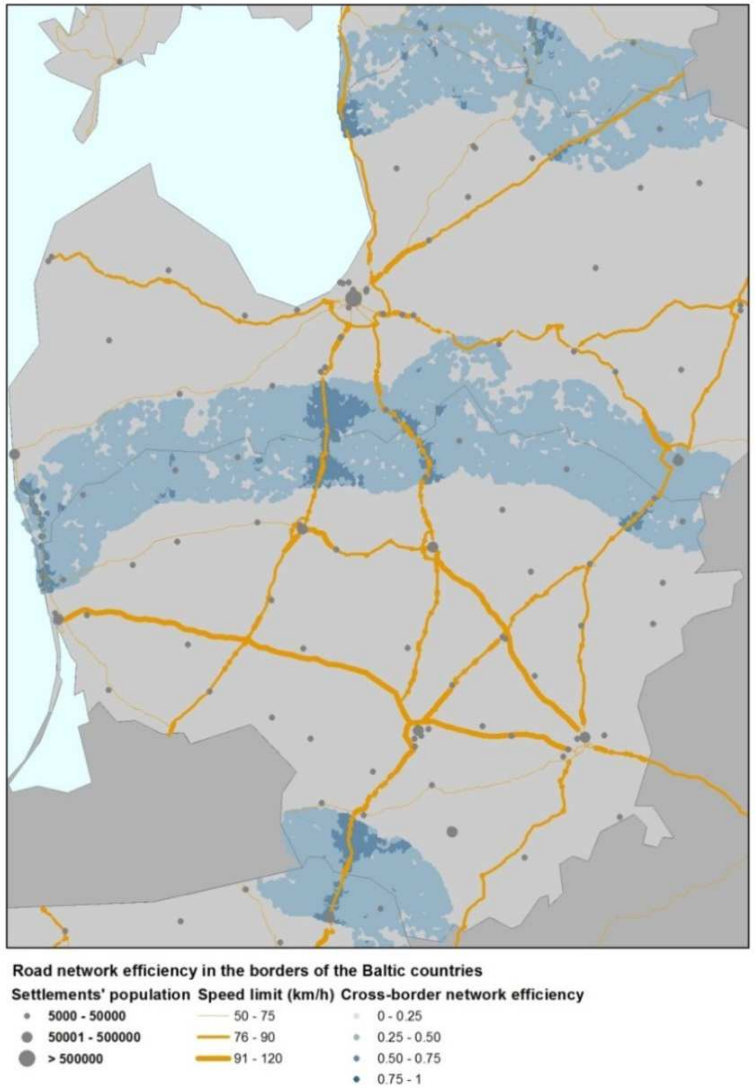
The first thing to observe in the maps with the results of network efficiency at grid level is the concentration of well performing points in close proximity to major roads leading to population settlements. This was expected, while it can also be seen that the distribution of the points with high network efficiency is affected by the location of main settlements.

The border zone between Latvia and Lithuania lies in the centre of the maps in Figure 25. The roads E77 and E67 are the ones connecting the LV-LT border region with Riga in Latvia and Šiauliai and Panevėžys in Lithuania respectively, and their impact on accessibility is obvious. Determining role in the western part of the border zone plays the connection to the city of Liepāja in Latvia as a result of the absence of other important settlements in the area.

Considering the LT-PL border, the concentration of the points with high network efficiency indicates the importance of E67 and E28 for the connection with Kaunas and Alytus respectively.

Figure 25. National (left) and cross-border (right) network efficiency in the Baltic countries





8.2 Border regions in the Balkans

A case with particularly low network efficiency is Bulgaria considering the borders with both Greece and Romania. The results of network efficiency for the borders between Greece, Bulgaria and Romania are presented in Figure 27.

At first it can be noted that national network efficiency is in general higher than cross-border, especially in Greece. This is not irrelevant to the fact that the vast majority of major settlements in Bulgaria are far from the borders, especially in the southern part of the country. The poor road network connection with Greece is obvious from the map of the major roads as there are practically only two main points to cross the GR-BG borders and very few other major roads within a buffer significantly larger than the 25km buffer determining the border area.

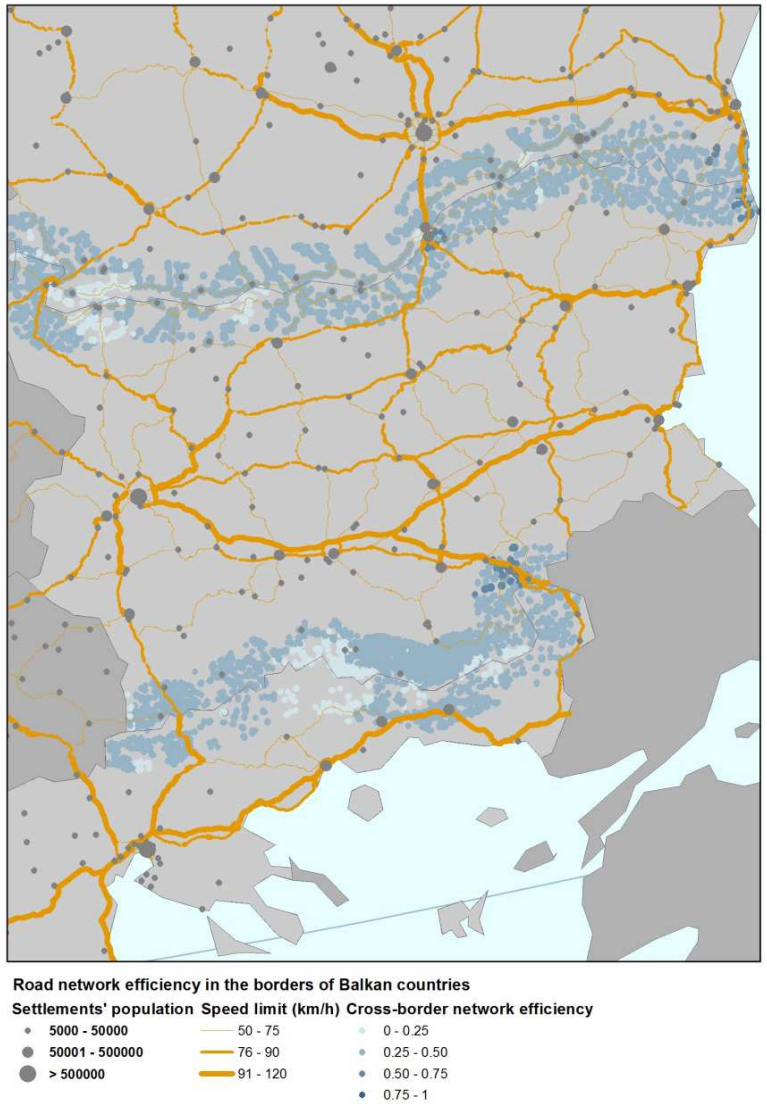
In both the border zones with Romania and Greece, there are very few grids with network efficiency larger than 0.5 in Bulgaria and they are concentrated around the road connecting the eastern part of Greece with Sophia and the northern part of Bulgaria with Bucharest or Burgas.

The borders of Bulgaria with Romania and Greece have quite particular morphology and this definitely affects accessibility. The borders between Greece and Bulgaria are on a mountainous and very little developed area, while the natural border between Bulgaria and Romania is the river Danube.

For the case of cross-border network efficiency there is high concentration of grids with network efficiency below 0.25 in the central part of the GR-BG border zone and the Western part of the RO-BG border zone. Finally, as can be seen in Figure 26 the border regions of Bulgaria with its EU neighbours are quite sparsely populated. The case will be discussed with even more detail in section 9.

Figure 26. Population at grid level in the area of Western Balkans

Figure 27. National (left) and cross-border (right) network efficiency in the Eastern Balkans



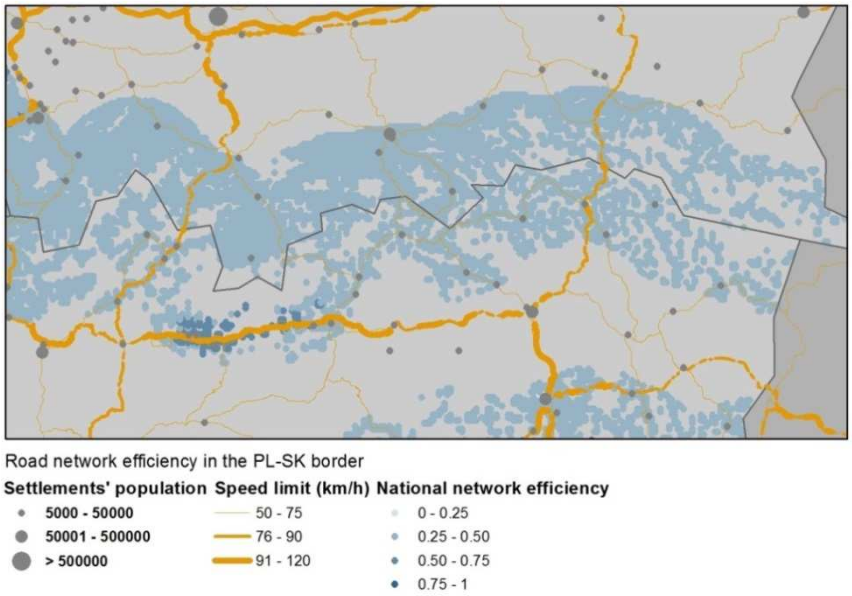
8.3 Border region between Poland and Slovakia

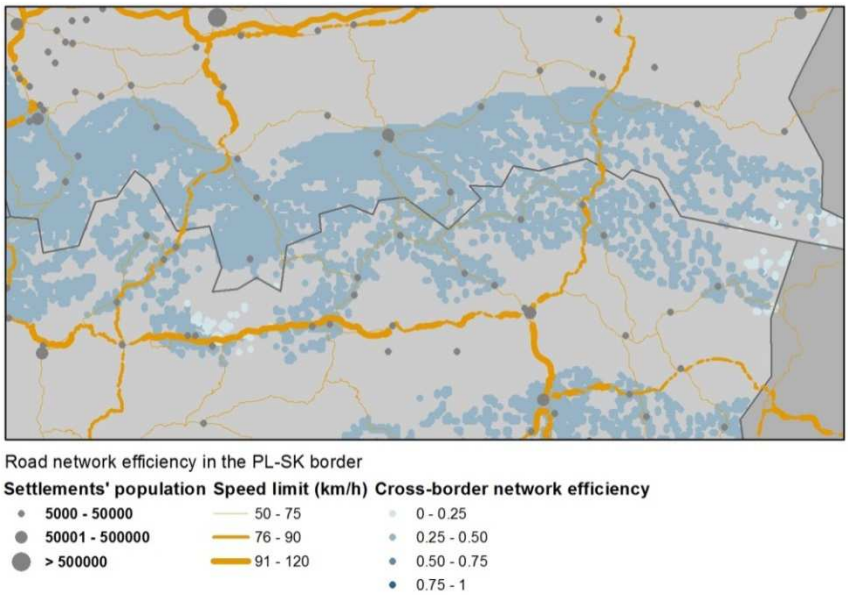
Another case of low network efficiency is the border zone between Poland and Slovakia. As can be seen in Figure 15, cross-border network efficiencies in both the Slovak and Polish parts of the zone are in the lowest band. Figure 28 presents the results of national and cross-border network efficiency.

Low cross-border network efficiency in the zone is driven by poor cross-border connections of the eastern part of the zone and by several low cross-border efficiency points in the Žilina region of Slovakia. As can be seen in both cases there are very few settlements in the areas and there are no major roads connecting them to the other side of the border. Both areas are close to or within the borders of national parks.

On the other hand, national network efficiency in the Žilina region is higher than in the rest of the border zone affected mainly by the existence of the E50 road.

Figure 28. National (upper map) and cross-border (lower map) network efficiency in the borders between Poland and Slovakia





8.4 Border region between Denmark and Sweden

The analysis at such a disaggregate level is very useful to examine the impacts of major infrastructure projects. A particularly interesting case is the Øresund bridge linking Denmark with Sweden. In Figure 29 the results of national and cross-border network efficiency are presented. As expected, the bridge has very high impact on cross-border efficiency. This happens in an area with relatively high population as shown in Figure 30, but of course this tunnel acts beneficially also for areas far beyond the borders of the two countries.

Figure 29. National (left) and cross-border (right) network efficiency in the Danish-Swedish borders

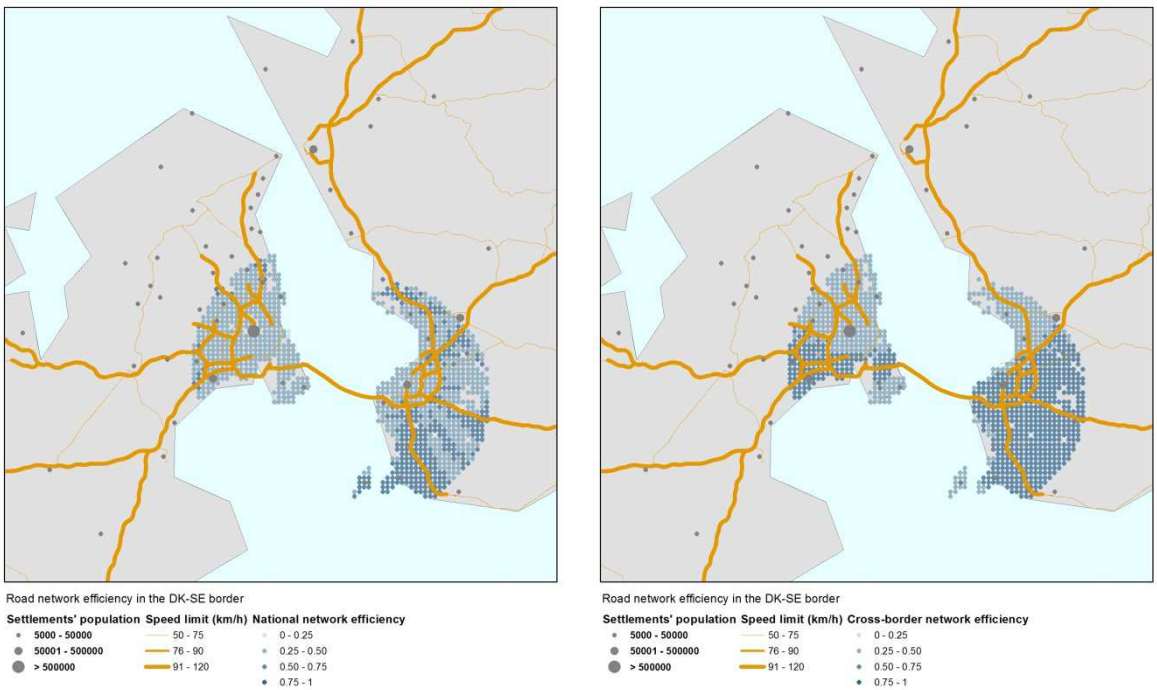
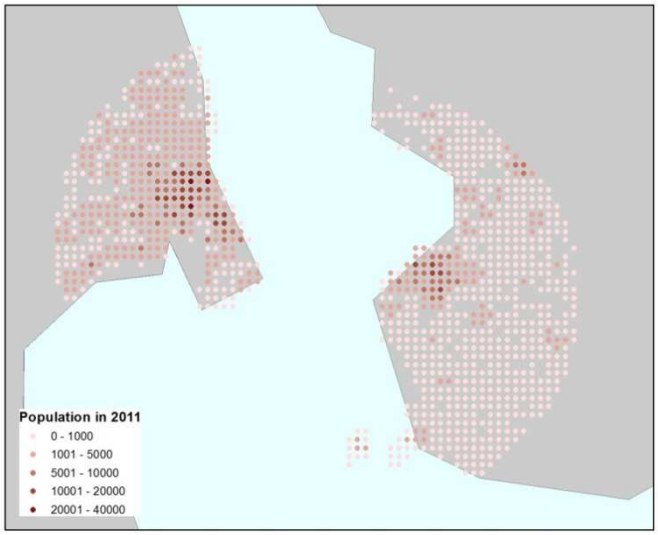


Figure 30. Population at grid level in the Danish-Swedish borders



8.5 Border region between France and Spain

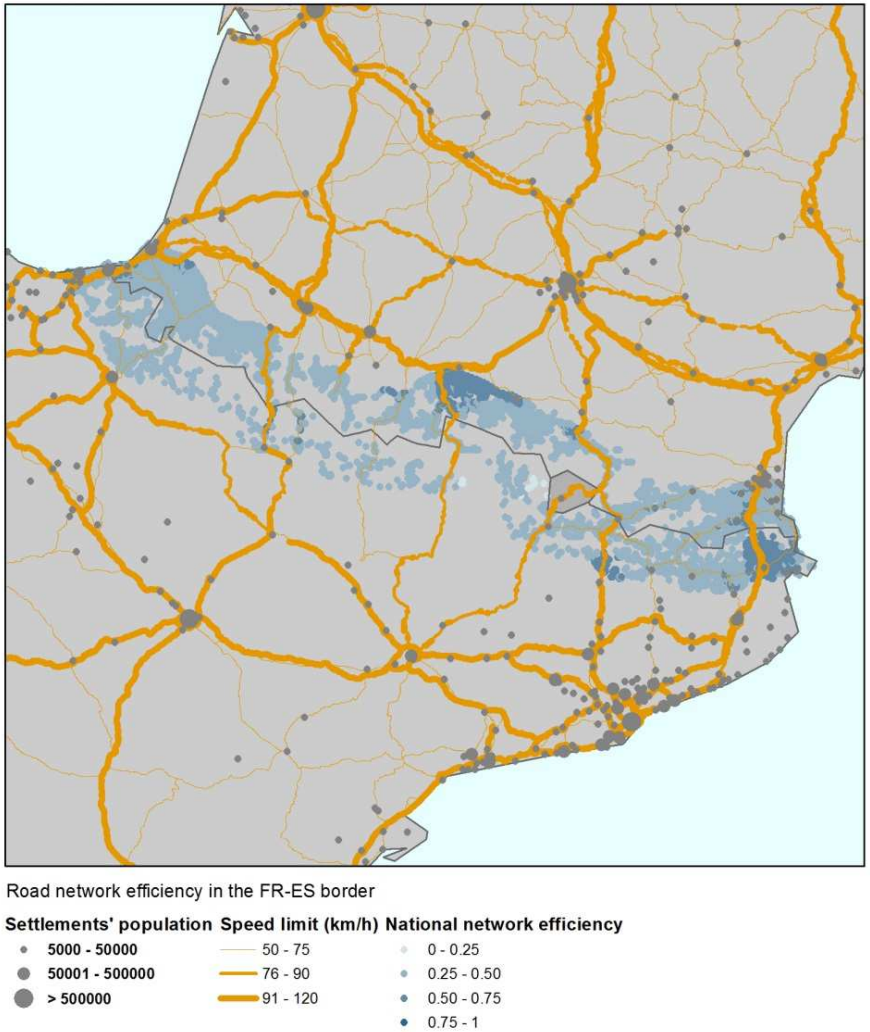
The aggregate results for the border zone between Spain and France (first column in Figure 15) show high cross-border network efficiency. From the results of national and cross-border network efficiency at grid level (Figure 31) it can be inferred that the aggregate results are mainly affected by the eastern and western parts of the border zone. These are the most populated parts of the zone and the ones with the best links across the borders. The concentration of population, settlements and major roads in specific parts of the border zone does explain the fact that although road network density is relatively low (Figure 22) network efficiency is high.

Although average cross-border network efficiency is high, there are several points with network efficiency below 0.25. Network efficiency at these points does not affect the aggregate results because the population there is relatively low.

An area with quite high national network efficiency in the centre of the French part of the border zone does not appear at all in the map with the cross-border results. The reason is that there are no settlements for these points within the determined buffer.

The best performing area considering both national and international network efficiency is the Spanish part of the eastern part of the border zone.

Figure 31. National (left) and cross-border (right) network efficiency in the French-Spanish borders



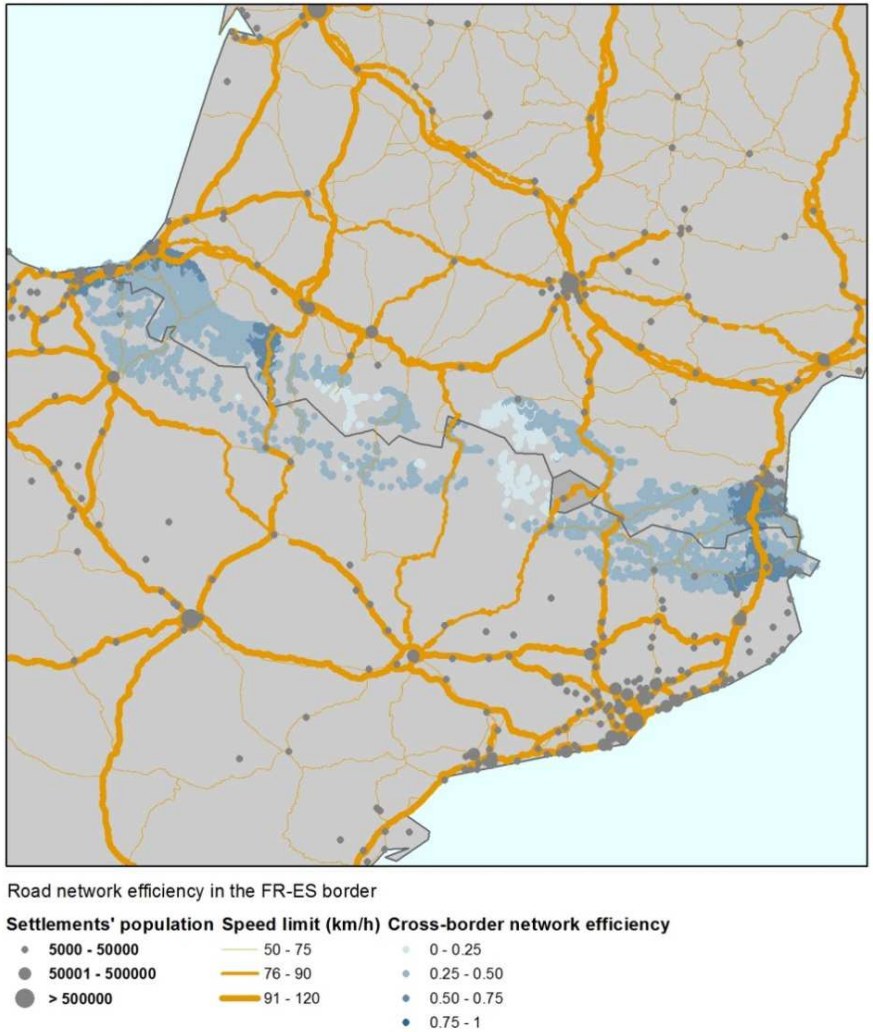
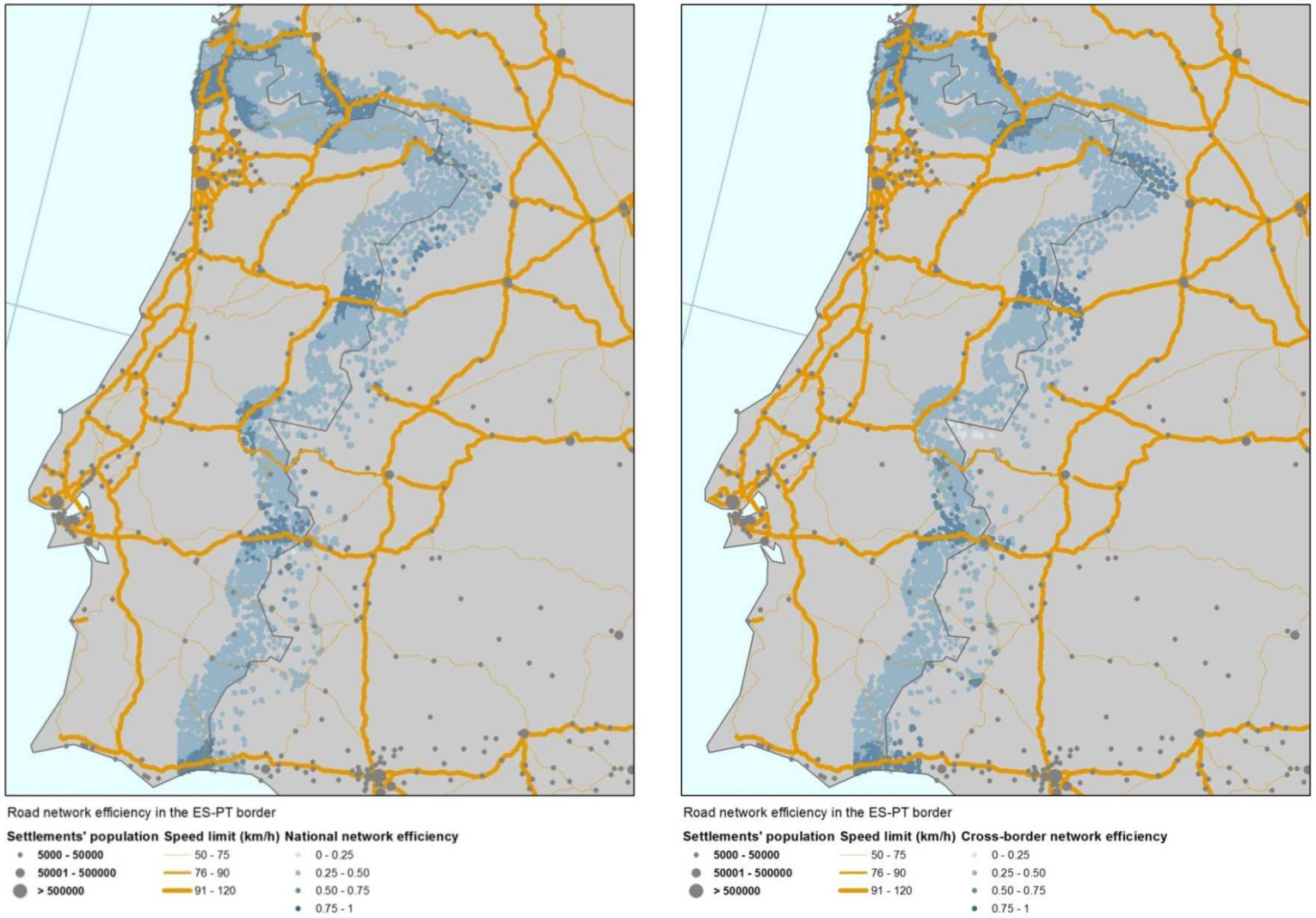


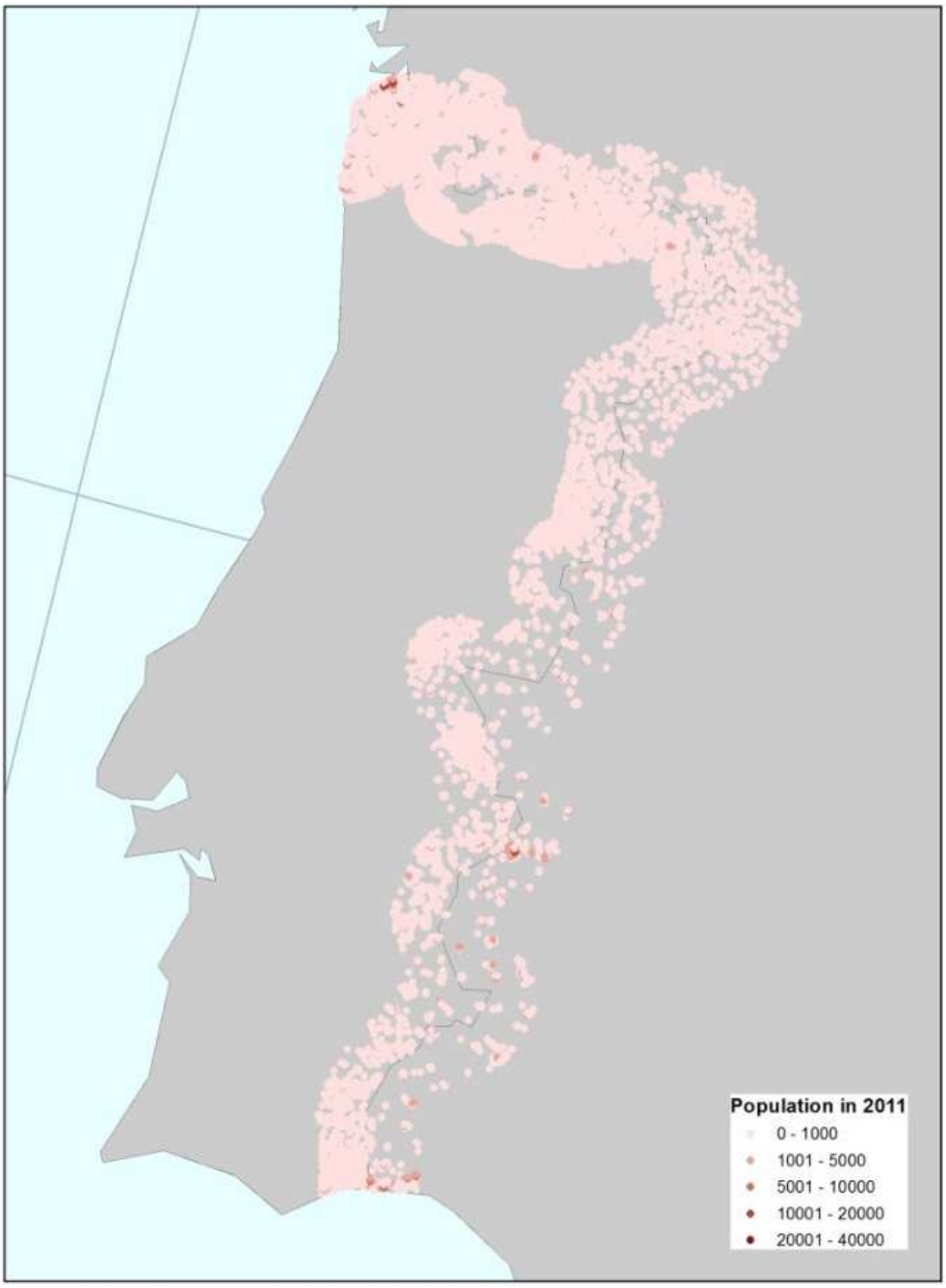
Figure 32. National (left) and cross-border (right) network efficiency in the Portuguese-Spanish borders



8.6 Border region between Spain and Portugal

In the border zone between Spain and Portugal (Figure 32) all major roads crossing the borders are surrounded by points with high network efficiency. The Portuguese part of the border zone is performing well in terms of both national and cross-border network efficiency. There are only a few major settlements in either side of the border and the most densely populated areas (Figure 33) are in the wider areas of major roads. As a result the border zone has high network efficiency but low network density.

Figure 33. Population at grid level in the Portuguese-Spanish borders



8.7 Border region between Germany and Poland

Regarding the border zone between Germany and Poland, cross-border efficiency for both countries is relatively high (Figure 15). For both Poland and Germany the aggregate results are driven by the fact that densely populated areas (Figure 34) have high cross-border network efficiency (Figure 35).

Figure 34. Population at grid level in the German-Polish borders

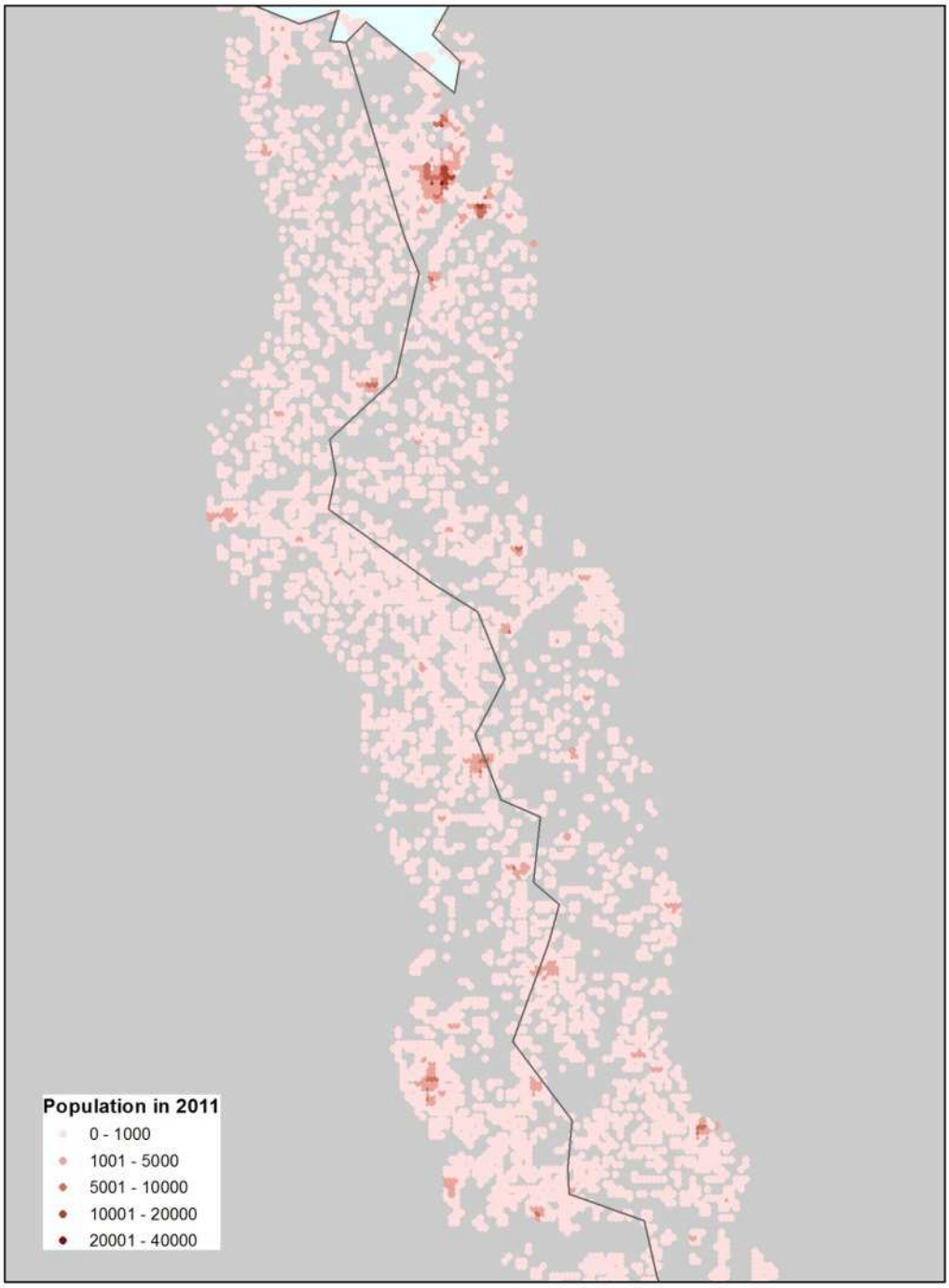
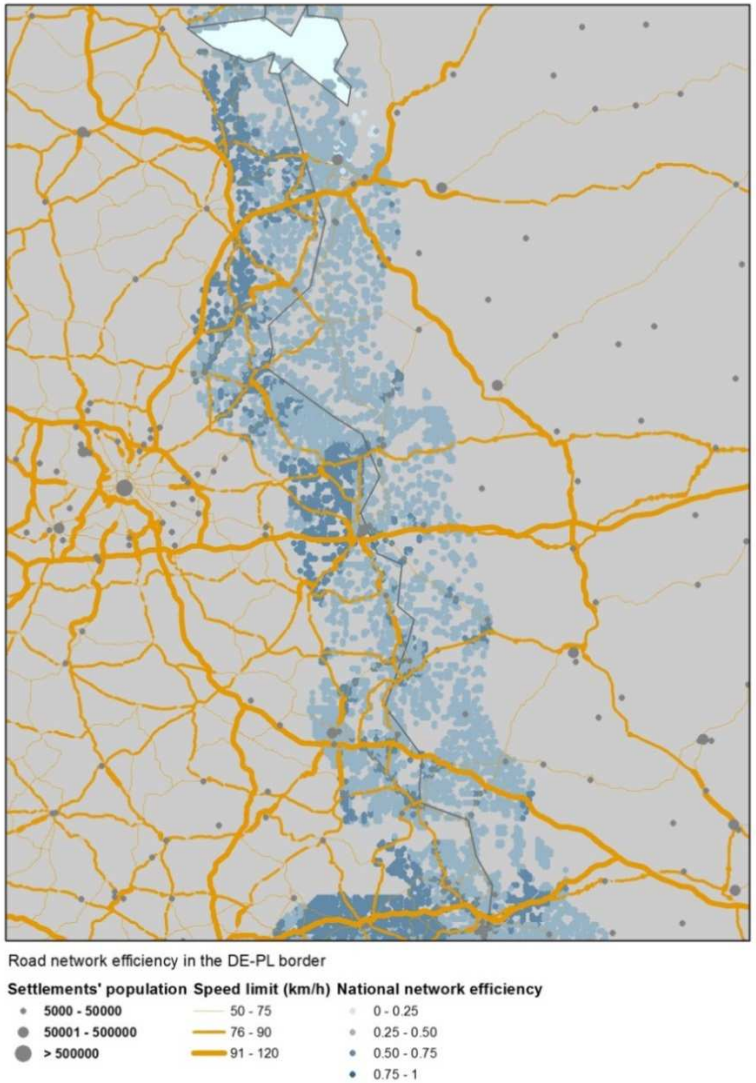
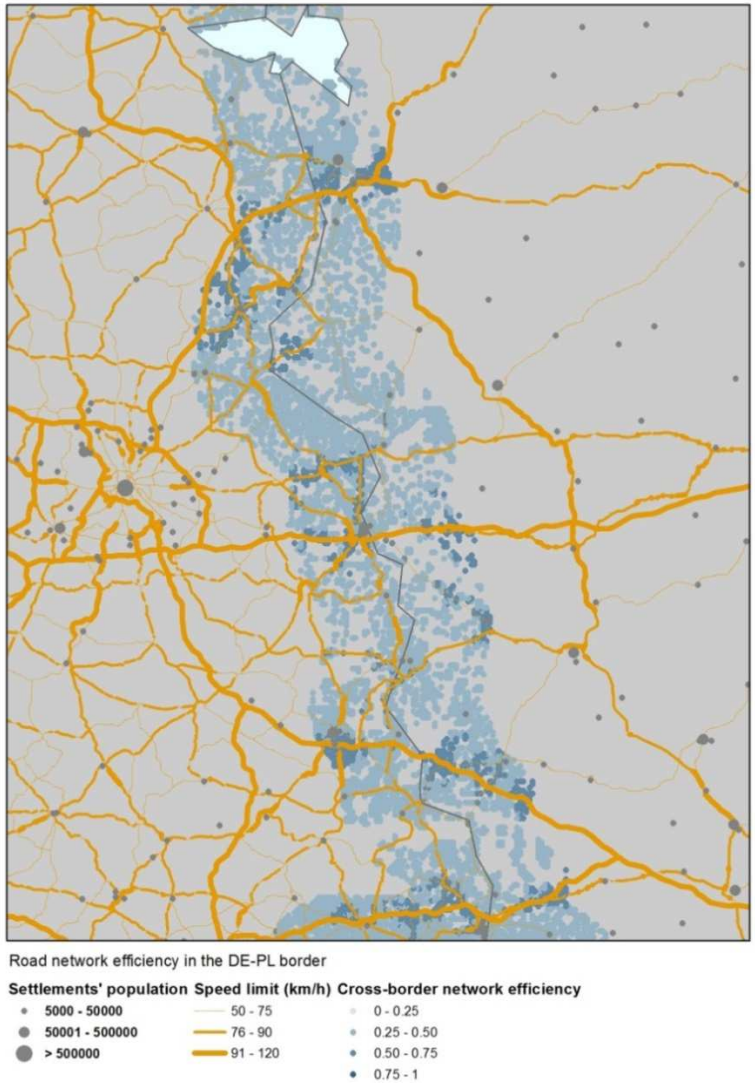


Figure 35. National (left) and cross-border (right) network efficiency in the German-Polish borders





9 Selecting cases using multiple criteria

By combining information on grid population (origins), settlements' population (destinations) and network efficiency it is possible to identify highly populated areas that lag behind in terms of connectivity and indicate potential infrastructure needs. All the available information has been brought together in a database that allows the selection of grid points based on a variety of criteria.

Figure 36 presents grid points selected according to the following criteria:

- Origin or grid population larger than the 0.95 quantile, corresponding to grids with more than 1210 people
- Destinations' population (total population of all the settlements selected as destinations from the specific grid) larger than the 0.5 quantile, corresponding to population higher than 218000 people
- Cross-border network efficiency lower than 0.4

These criteria were determined in order to identify grid cells with high population located close (within the determined buffer) to settlements with relatively high population but with low cross-border network efficiency. These could be cases where infrastructure investments could return relatively high benefits.

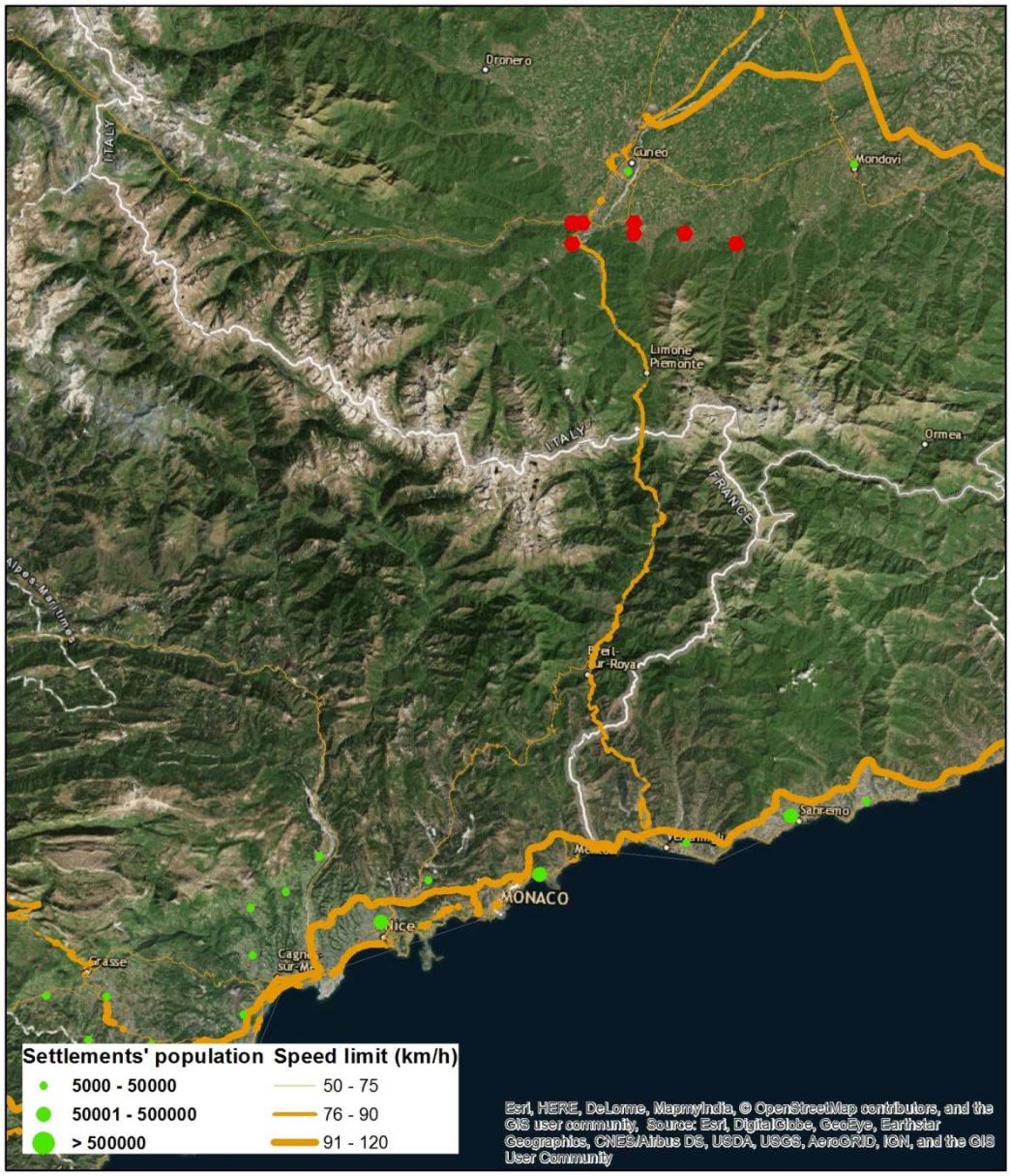
A selection of cases indicated in Figure 36 will be discussed in more detail in this section.

Figure 36. Selection of grid cells with relatively high origin and destination population and low network efficiency

9.1 Selected area in the border zone between Italy and France

The first area to zoom in is the group of grid cells in the French-Italian border zone (Figure 37). In total 7 points have been selected amounting more than 16000 people. All the destination settlements are on or close to the French coast and there is only one major road connecting the area with the coast. The area is sparsely populated and very mountainous which explains the fact that the road network is not very developed. Natural borders and demand are determining factors of the quality and density of infrastructure.

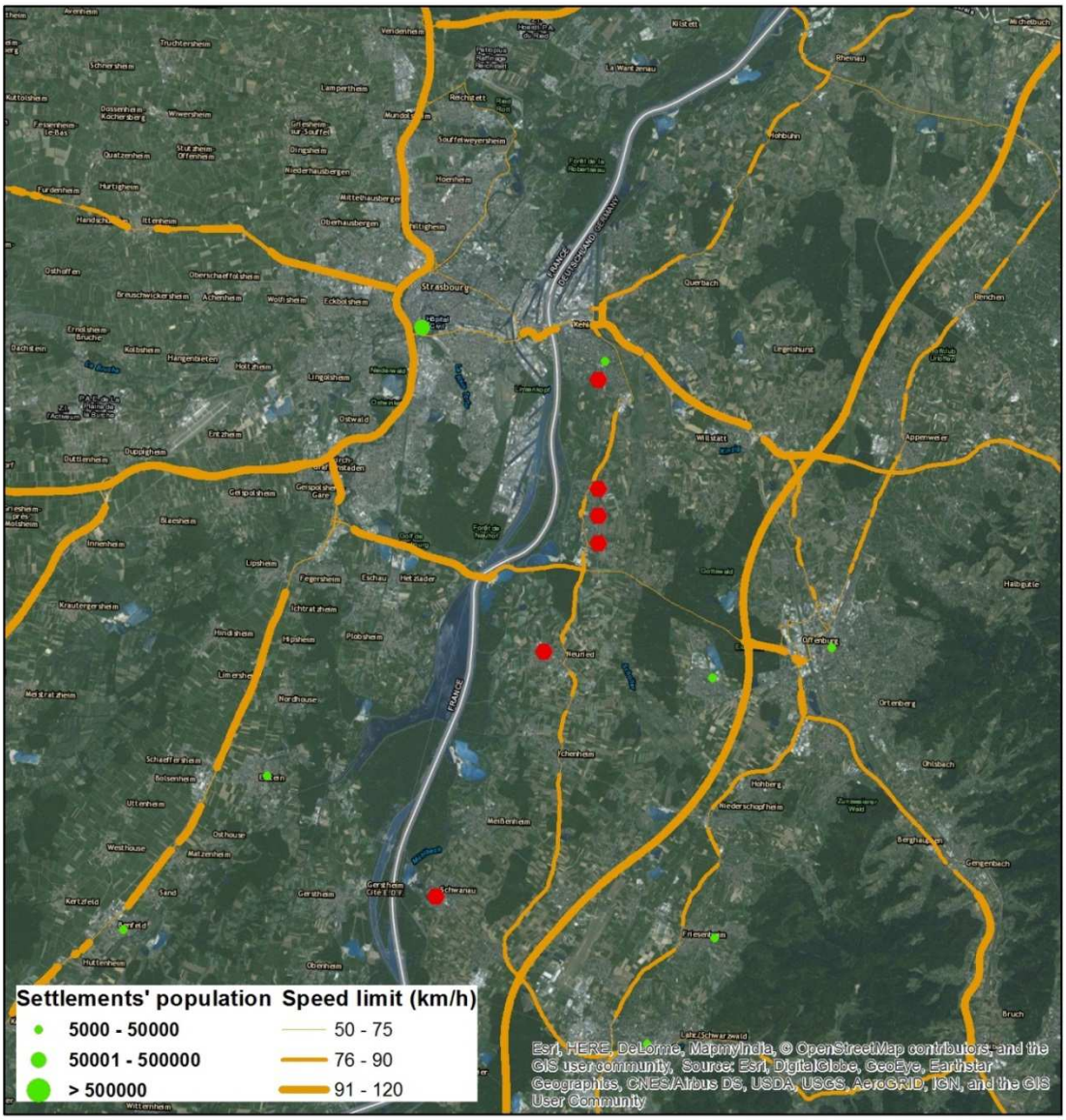
Figure 37. Selected grid cells close to the borders of Italy with France



9.2 Selected area in the border zone between Germany and France

Another case where natural borders play important role is the border between France and Germany in the area between Strasburg in France and Schwanau in Germany. This time the natural border is river Rhine as can be seen in Figure 38. There are two major roads connecting the two parts of the borders and the connection to Strasburg, which is the biggest city in the area, plays a determining role in the results of network efficiency in the area.

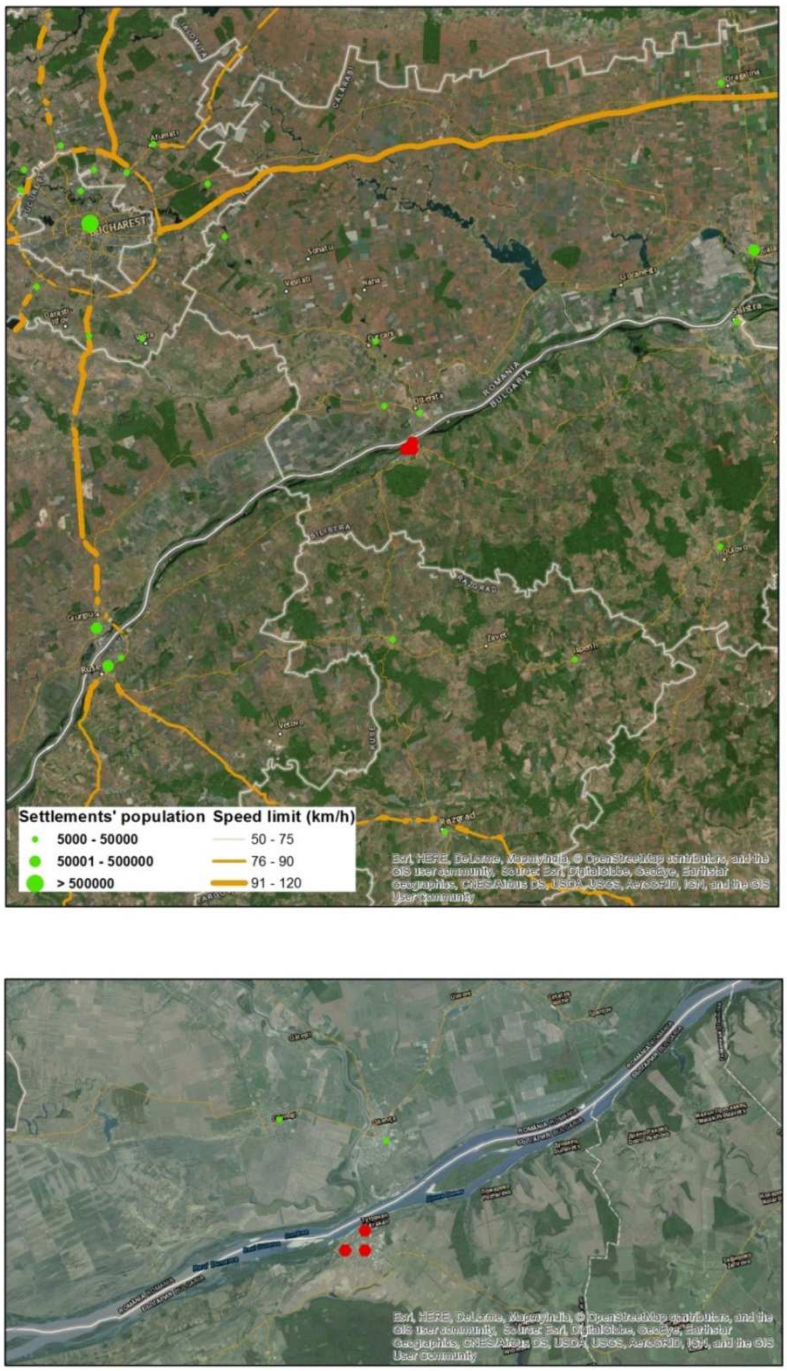
Figure 38. Selected grid cells close to the borders of Germany with France



9.3 Selected area in the border zone between Bulgaria and Romania

The existence of a river as a natural border, Danube this time, determines also network efficiency measurements of the border zone between Bulgaria and Romania. In this case, there are only three grid cells selected with more than 6000 people in total but the case highlights the effects of the scarcity of bridges in the area. As navigation is not considered in the analysis, for the selected points in Bulgaria somebody would have to make a 135km road trip in order to reach a relatively large settlement (with more than 5000 people) in Romania that is less than 5km away (straight line distance).

Figure 39. Selected grid cells close to the borders of Bulgaria with Romania (zoom in view in lower map)



9.4 Selected areas in the border zone between Czech Republic and Poland

The highest concentration of selected points (according to the criteria determined at the beginning of this section) is found on the borders between Czech Republic and Poland (Figure 40). To identify particularly problematic cases an additional parameter is used. In addition to the selection criteria used until now, national network efficiency is also considered. Hence, highly populated grid cells that are poorly connected to relatively large settlements in both the country where they are located and the neighbouring one are selected. They are marked with black points in Figure 41, while with green are marked the settlements considered to measure network efficiency.

Figure 40. Selected grid cells close to the borders of Czech Republic with Poland

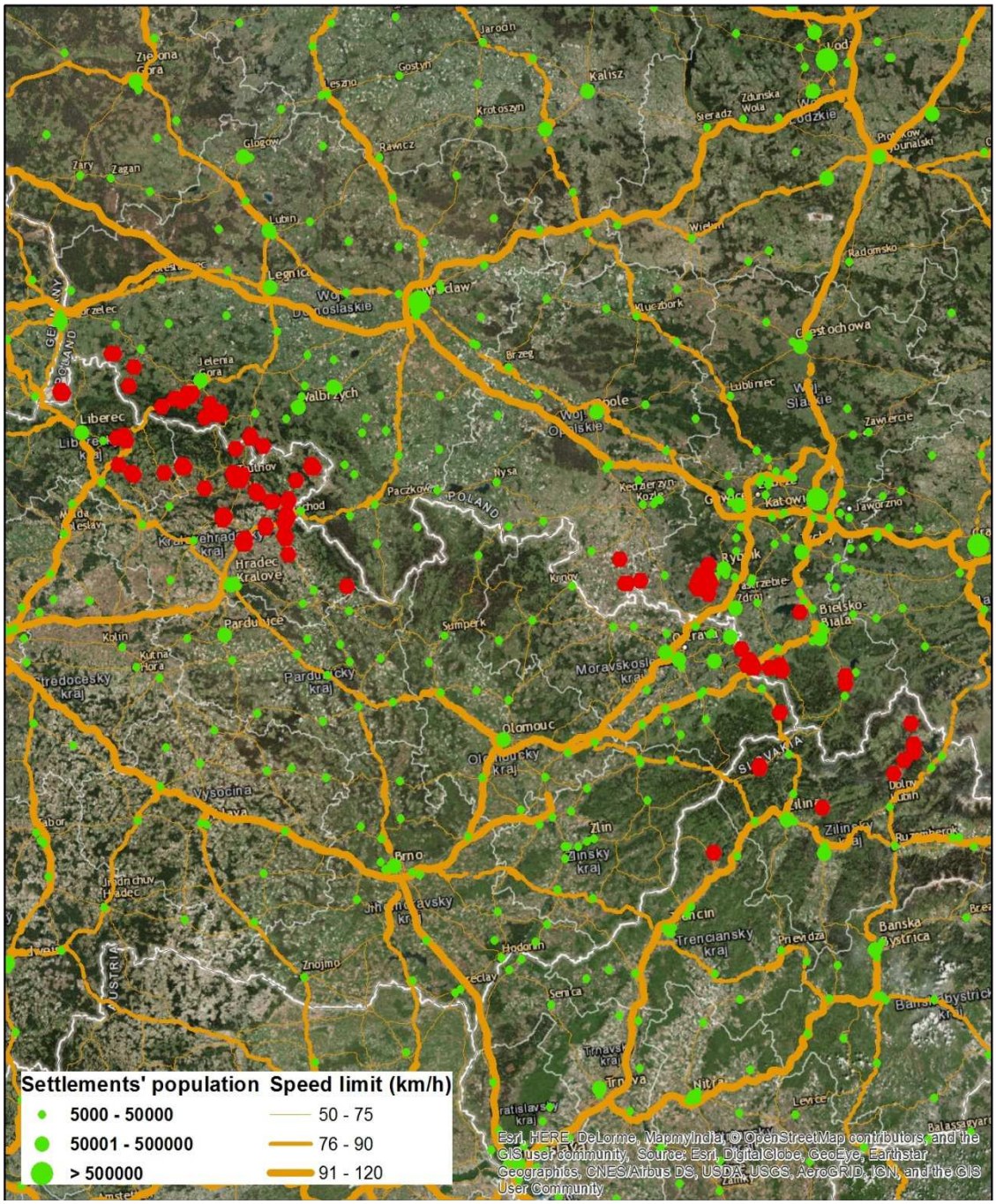
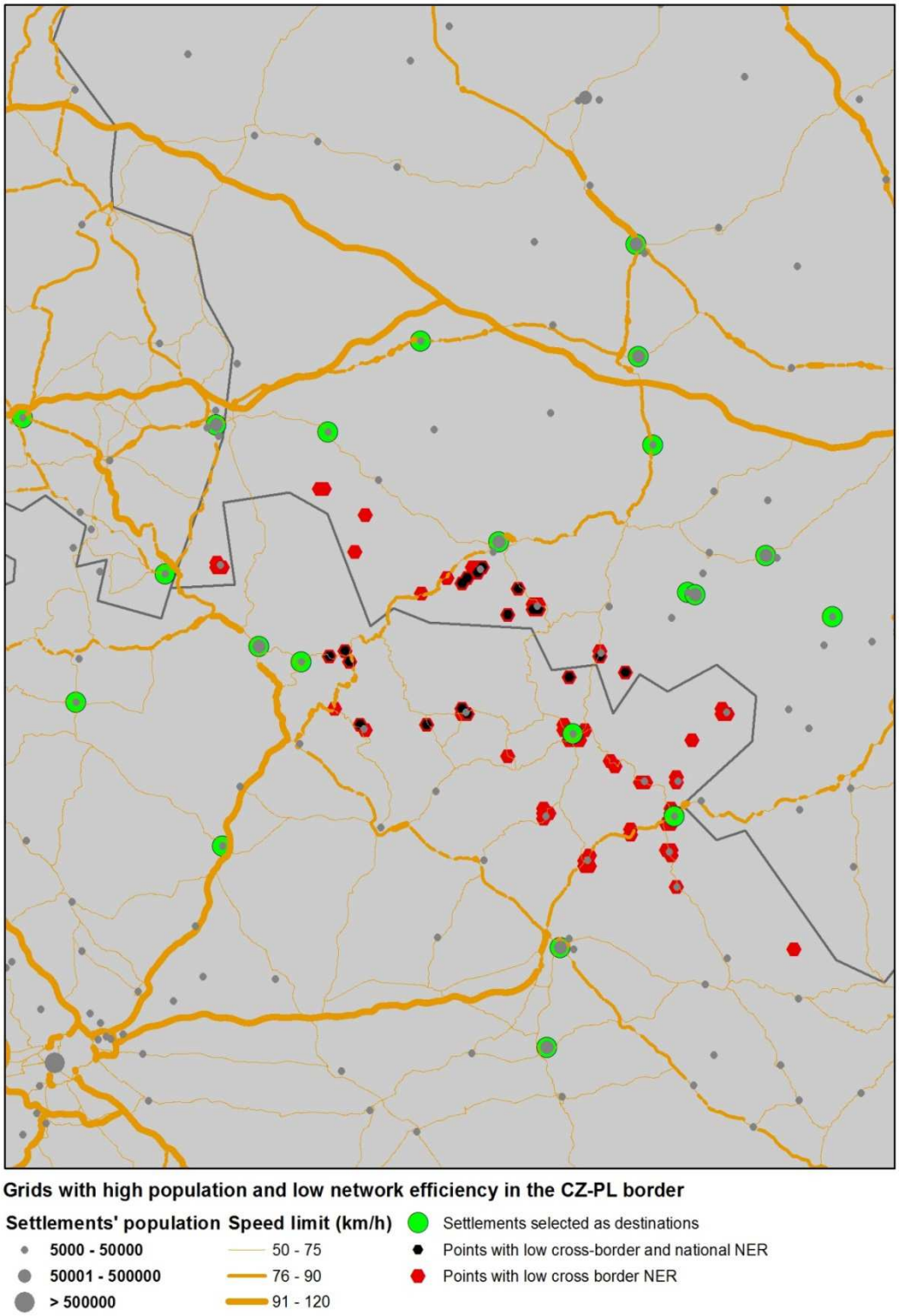


Figure 41. Selected grid cells close to the borders of Czech Republic with Poland considering both national and cross-border network efficiency



Having identified potentially problematic cases as the ones already discussed, more detailed analysis would be required for each case separately in order to understand whether inefficiencies have to do with design, speed etc. and to improve the level of information required for decision making. This would mean to analyse specific couples of origins (grid cells) and destinations (settlements) focusing on specific routes. The tools developed in the context of this study, make this type of analysis feasible as it is possible to extract and map the specific routes of each shortest path calculated.

10 Application for policy analysis

The methodology presented in the preceding sections allows a broad range of analyses related to border regions and/or cross-border collaboration and can be potentially extended to the exploration of regional development issues in general. While the approach is based on broad accessibility concepts, the indicators used combined with the specific geographic focus selected for this particular implementation allow the generation of relevant quantitative input to policy processes, especially to the ones aiming to:

- Ensure access to a defined level of basic or specific services
- Increase opportunities and/or economic activity
- Improve quality of road transport network
- Decrease isolation/ remoteness

The approach and the results are obviously only a tool for the policy analyst, they do not consist a policy optimization model. As the step-by-step description of the approach suggests, a number of the parameters that affect the definition of the interest areas and the operationalisation of the indicators are selected empirically by either the policy analysts or the application developers. The translation of policy-side definitions into numerical values involves many approximations and leads to inevitable uncertainty and imprecision. Nevertheless, if those limitations are taken into account, the methodology provides a useful tool that helps quantify and visualise otherwise challenging concepts.

The first step of the approach consists of the selection of parameters to reflect the policy side definitions:

- Level of geospatial detail: 1 km x 1 km grid
- Border zone definitions: 25 km from land border
- Area of interest around each zone: radius of 75 km from each grid cell's centre
- Number and size of settlements accessed: 5 settlements with population over 5 thousand

This combination of parameters has been selected empirically, through a close collaboration between the policy analysts and the application developers, including a number of tests using different thresholds. All choices affect the level of relevance and precision of the results and create a trade-off with computational requirements. Those thresholds appear to provide a detailed enough granulation and a rational approximation of what can be considered a border zone and what is relevant in terms of cross-border interactions.

The second step consists of selecting the indicators to be calculated. Three indicators were selected in this exercise, chosen based on how the policy objective was defined:

- Location indicator
- Potential accessibility
- Network efficiency

Additional or different indicators can be combined in this approach depending on the policy focus. Even for the specific set selected, there are additional options as regards their operationalisation and the parameters used, the most important ones being the attraction variable (in this case population), the decay function (linear in the case of Potential Accessibility) and the distance calculation (A* shortest path calculation on the TomTom network).

10.1 Prioritisation using policy based criteria

The dataset used can be the basis for a filtering process that can be useful for the prioritisation of policy initiatives aiming to improve the cross-border collaboration across Europe. As already discussed, the design of policy initiatives on regional development and –more specifically- border area socio-economic issues is quite complicated since many cultural, political, historical, geographic, demographic and economic conditions affect the degree at which two cross-border zones interact today, as well as how this interaction can change as a result of improving transport infrastructure.

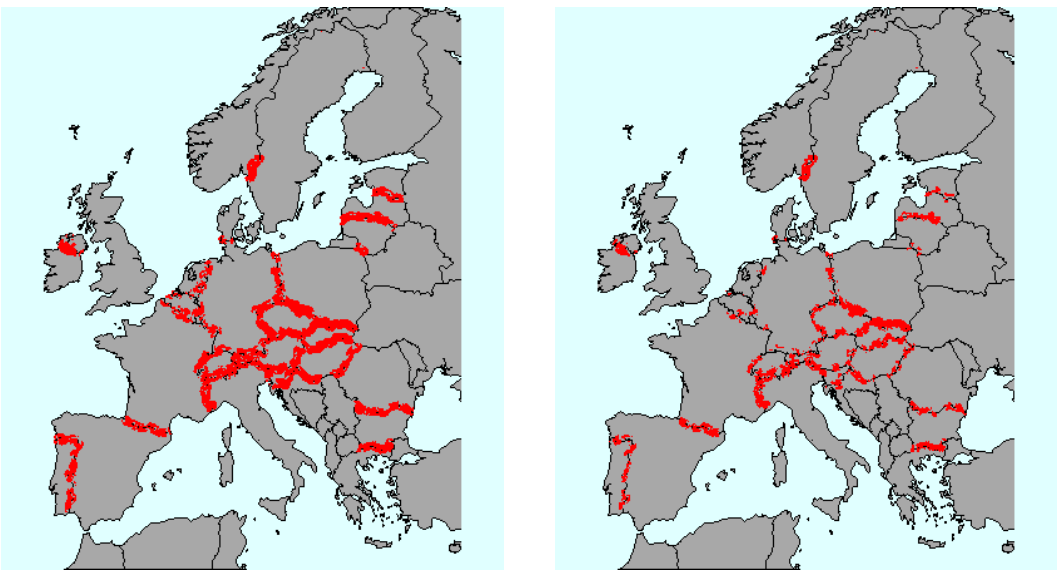
The dataset used here can however be useful in describing the existing situation and identifying areas where an improvement in transport infrastructure could potentially lead in improving one or more of the policy objectives set. The three indicators used allow flexibility in terms of defining the policy objective and evaluating how this is met. Identifying areas for policy intervention however is a heuristic rather than a determinist process, and does not have a clear recipe for reaching the perfect solution. Nevertheless, a combination of suitably worded policy objectives with a robust set of quantitative indicators can at least guide policy makers towards options that can approximate the aims of the policy objectives.

A few examples of how this methodology and the resulting dataset can be used in the context of specific objectives:

- **Increase access to basic cross-border services**
    a. Assume goal is that Location Indicator (international) should be lower than 1 (i.e. the average travel time to the top 5 cross-border settlements should be less than 1 hour)
    b. Filter dataset to include only zones with Location Indicator >1 (Figure 42 left)
    c. Filter dataset to include only zones with Location Indicator >1 to identify areas with long travel time (Location Indicator) also from the national side (Figure 42 right)

The resulting list of zones can be considered as potential candidates for measures aim at decreasing isolation through the improvements of cross-border road links. Nevertheless, an additional case-by-case analysis of the local situation is necessary in order to explore the feasibility and expected impact of such measures.

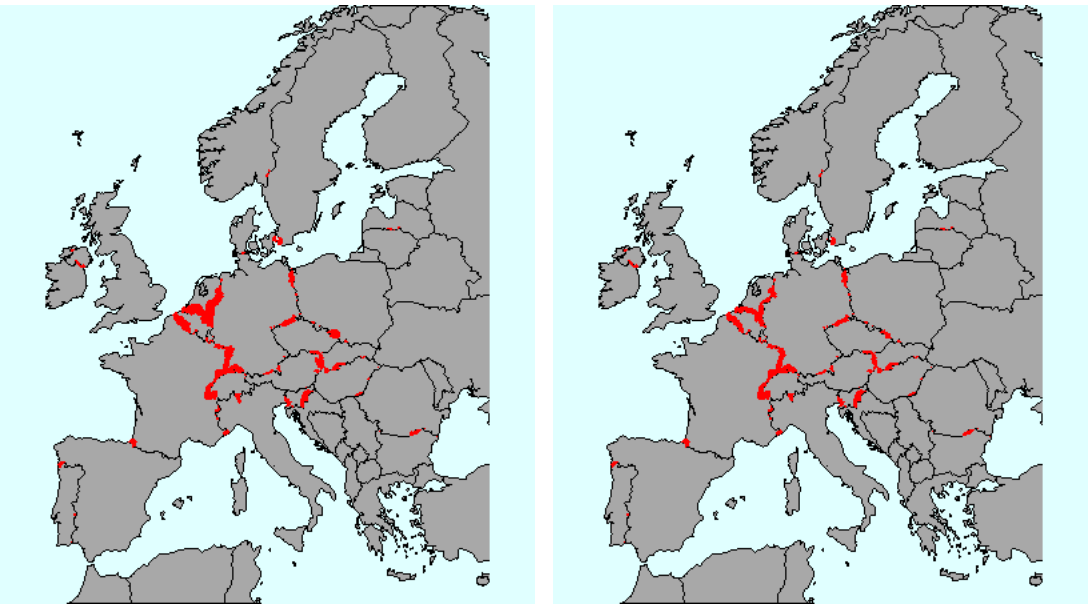
Figure 42. Using Location indicator to identify areas with limited access to services



- **Identify areas with cross-border opportunities**
    a. Filter dataset to include only zones with Potential indicator above certain threshold (0.5 million in the case of Figure 43 left)
    b. Use additional filter to identify zones where cross-border potential is higher than the national side potential (2x in the case of Figure 43 right)

A common assumption in regional analysis is that the Potential Indicator is a proxy for economic opportunities. Under this assumption, the identified zones have access to a cross-border population of significant size (1/2 million inhabitants within a 1-hour travel time) which can stimulate collaboration. The majority of those zones is near important urban centres and -predominantly- in the already most developed parts of the EU.

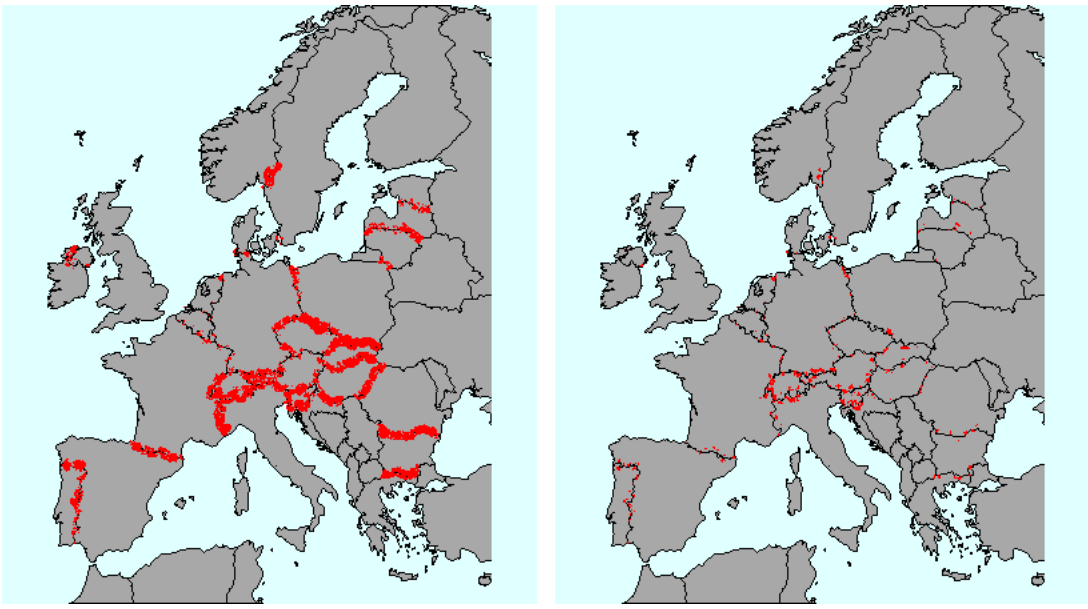
Figure 43. Using Potential Accessibility indicator to identify areas with cross-border economic opportunities



- **Identify areas with low network efficiency**
    a. Filter dataset to include only zones with Network Efficiency indicator below certain threshold (0.4 in the case of Figure 44 left)
    b. Use additional filter to limit to zones with good infrastructure quality on the national side, higher than 0.5 in the case Figure 44 (right):

The resulting subset of zones has low network cross-border efficiency while the national side network efficiency is relatively high. The reasons may be purely geographic, but this can also mean that there is opportunity for improvement on the cross-border side.

Figure 44. Using Network Efficiency indicator to identify areas with low transport infrastructure quality



- **Combining indicators to measure connectivity**

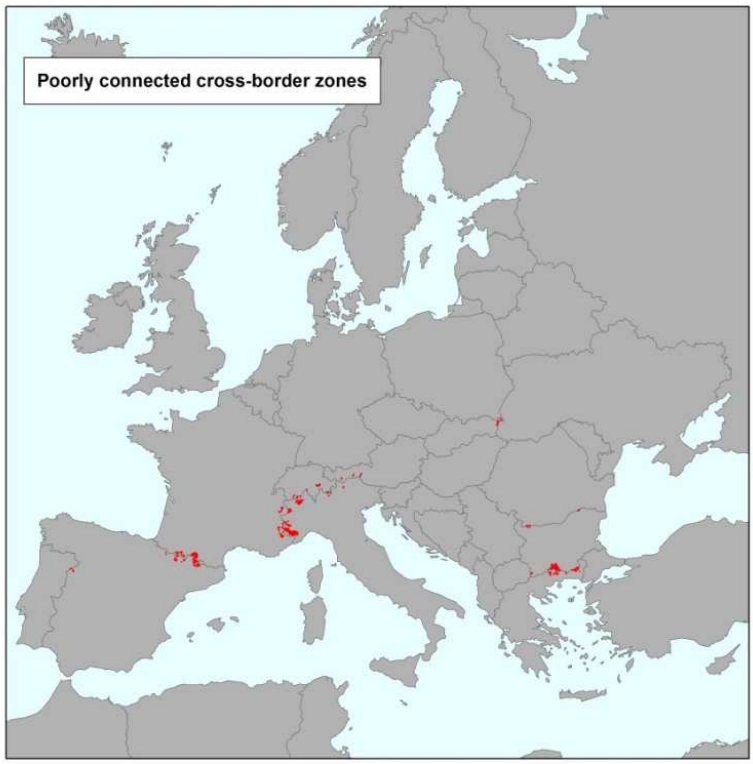
Using different combinations of indicators and thresholds allows the identification of border zones meeting several conditions. Three examples of filtering through multiple indicator thresholds provide grouping of policy interest from different perspectives:

a. Zones with low cross-border connectivity (Figure 45):

Zones with Location indicator (international) > 2 and

Network efficiency (international) < 0.3

Figure 45. Zones with low cross-border connectivity and accessibility



b. Border zones with high connectivity and accessibility on both sides (Figure 46):

Location indicator (international) < 1

Location indicator (national) < 1

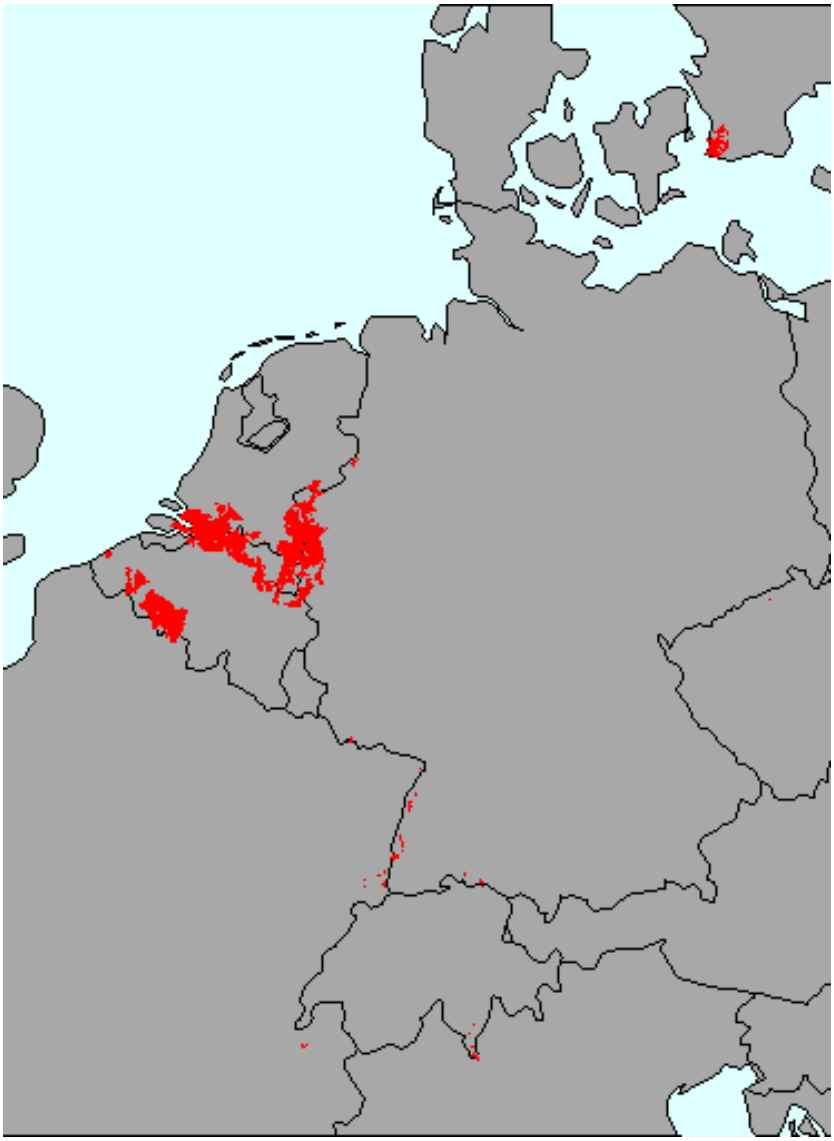
Network efficiency (international) > 0.5

Network efficiency (national) > 0.5

Potential accessibility (international) > 1 000 000

Potential accessibility (national) > 1 000 000

Figure 46. Zones with high connectivity and accessibility on both sides



c. Border zones with high connectivity and accessibility on international side (Figure 47):

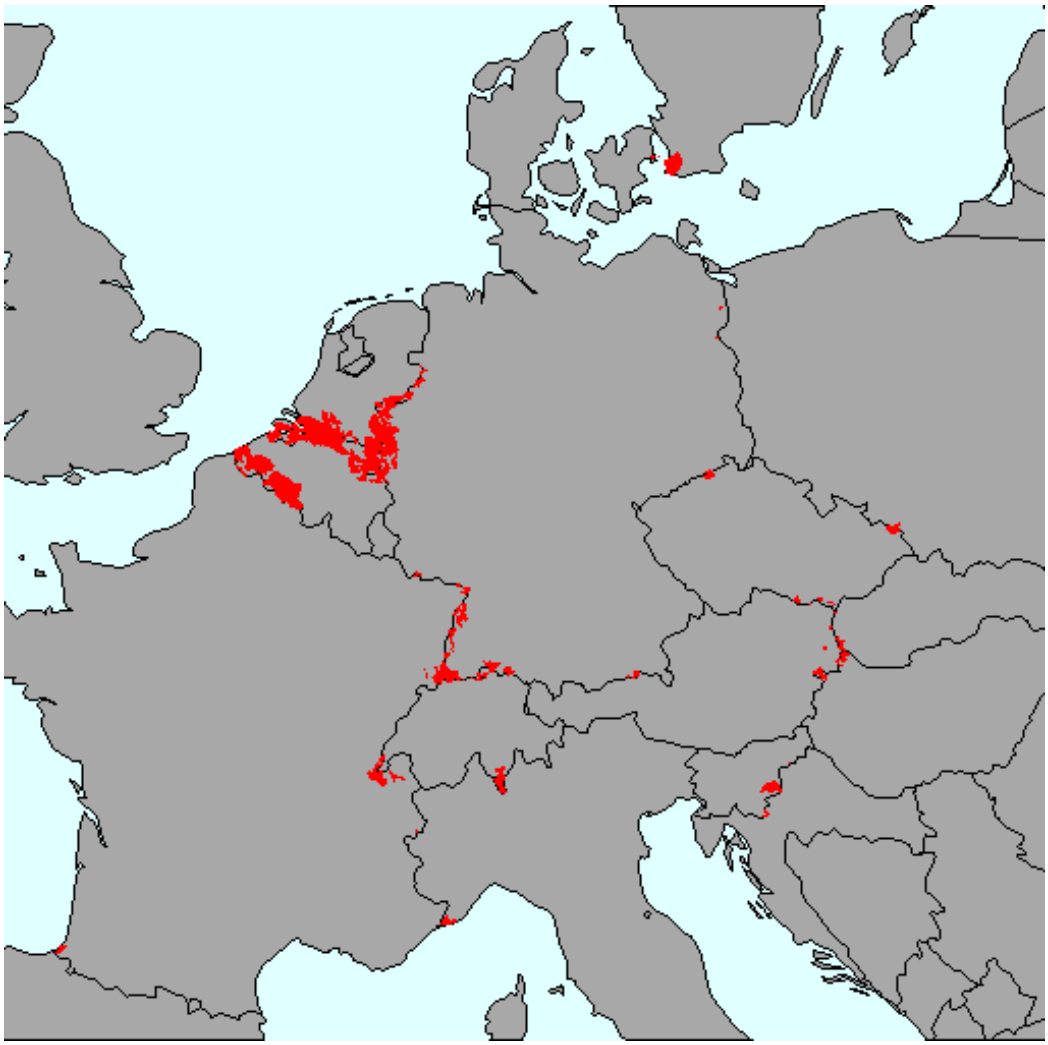
Same as previous example without restrictions on national side

Location indicator (international) < 1

Network efficiency (international) > 0.5

Potential accessibility (international) > 1 000 000

Figure 47. Zones with high cross-border connectivity and accessibility



- **Estimating impact from network improvement**

Improving cross-border road links can decrease travel time and improve network efficiency, but would also decrease the Location Indicator and increase the Potential Accessibility indicator. While the decrease in the Location Indicator is lineary correlated with the decrease in travel times, the correlation of the Potential Accessibility indicator with travel time is curvilinear (travel time is in the denominator). Exploring how the two indicators change as a result of improving network efficiency can give useful input to policy makers.

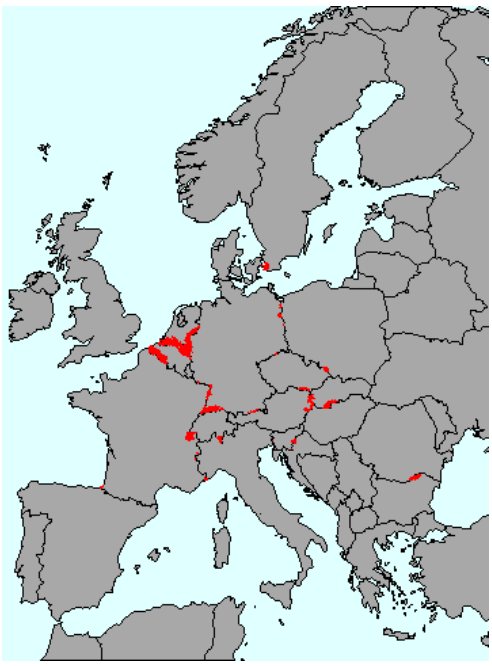
As a first example, the zones where the highest improvement in terms of Location Indicator as a result of improving network infrastructure can be identified. If the same percentage improvement is assumed for all zones in the dataset, the benefit as regards the reduction of the location indicator in absolute terms will differ, depending on the combinations of the actual values of the location and network efficiency indictators. Assuming a 10% improvement in network efficiency and using a 10 minute threshold value for decreases in the international location indicator produces a map of the areas where the impact is the highest (Figure 48).

In the second example, the impact in terms of increase in international potential accessibility is explored. A 10% increase in network efficiency is also assumed here, with the threhold of increase in potential accessibility, the size of the population accessed within 1 hour of travel time, set at 100 thousand (Figure 49).

The two results are drastically different, since two distinct policy objectives are assessed. In the first case, the highest benefit is expected in areas that actually combine low network efficiency and high location indicator (i.e. long travel times). In practical terms, this means that improving roads benefits most isolated areas the most. In the second case, the benefit clearly depends on the opportunities already available in the area. The zones already enjoying high accessibility would have a higher benefit, in absolute terms, than less accessible zones, for a comparable improvement in road networks.

Figure 48. Zones for which a 10% in network efficiency results in a reduction of the international Location Indicator by more than 10 minutes

Figure 49. Zones for which a 10% in network efficiency results in an increase of the international Potential Accessibility Indicator by more than 100 thousand (accessible population per hour of travel time)



The difference in the profiles of zones identified through the use of each indicator is even more visible when they are compared on a correlation graph. Figure 50 shows the correlation of the location indicator with network efficiency, both at cross-border level. The red dots represent the areas identified as having the highest impact in terms of location indicator improvements (and mapped in Figure 48), while the red dots are the ones where potential accessibility increases the most (mapped in Figure 49). A similar comparison is made for the correlation between potential accessibility and network efficiency (again at cross-border level for both) in Figure 51.

Figure 50. Correlation of location indicator and network efficiency, cross-border (red= zones in Figure 48, green= zones in Figure 49)

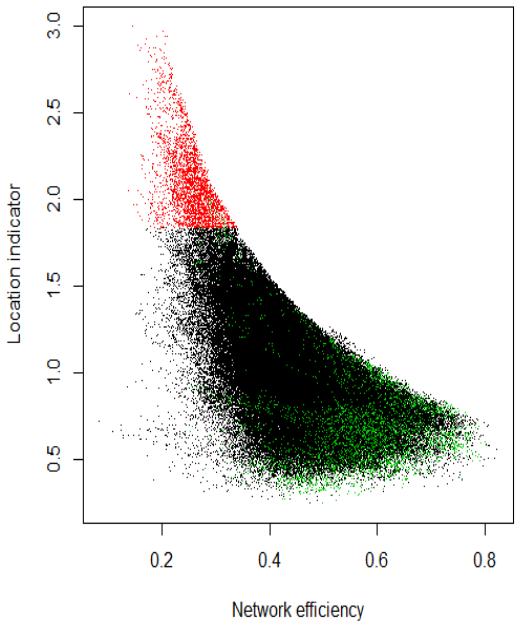
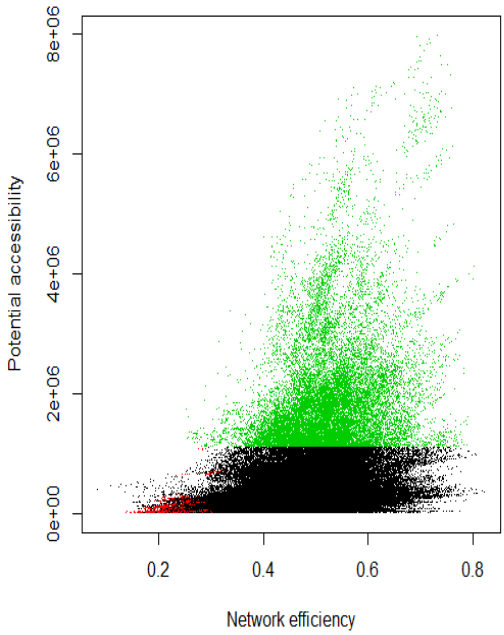


Figure 51. Correlation of potential accessibility and network efficiency, cross-border (red= zones in Figure 48, green= zones in Figure 49)



10.2 Machine learning techniques

An extension of the methodology described here addresses option to automate the process of identifying areas of policy interest by using machine learning techniques to process the data and implement different policy criteria. Given the availability of extensive quantitative data, either as a result of the calculation of the indicators discussed here or in combination with data from external sources, several additional types of input to policy makers are possible. Since however -as already discussed- many of the policy priorities in regional development entail qualitative elements that require empirical input and/or expert-based heuristics, it is not possible to identify an *optimum* solution that clearly identifies where to focus policy interventions.

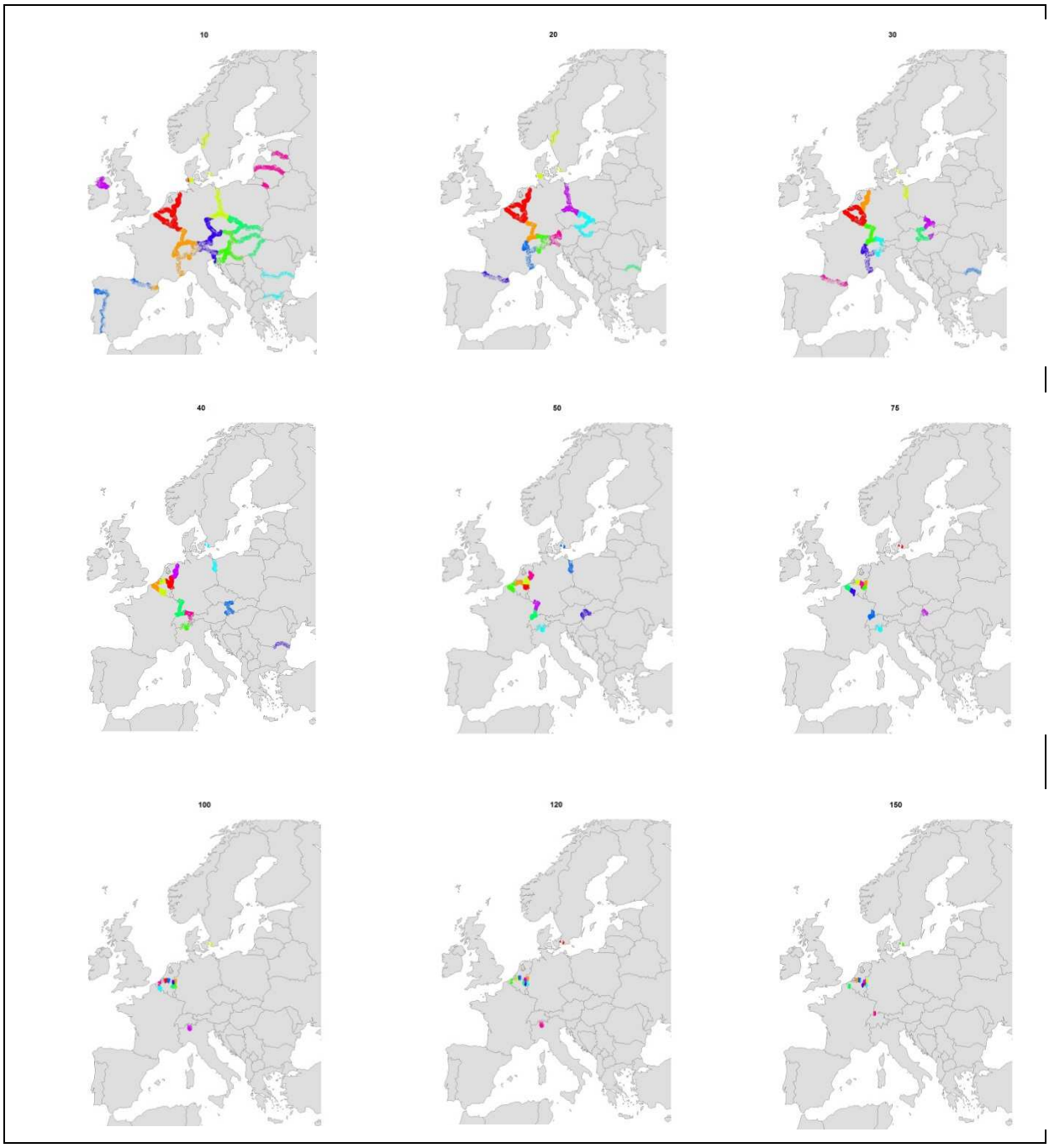
Machine learning techniques and –in this particular case- clustering techniques allow additional experimentation and analysis when addressing an open ended question such as the identification of areas with maximum potential benefit. An iterative approach based on k-means clustering allows the calculation of potential benefits (under any user-defined definition of benefits) for varying combinations of neighbouring zones. Such an analysis can aid policy makers in prioritizing investments by area of interest using different size criteria for the definition of each are.

The k-means clustering approach consists of dividing data into a number of k clusters with the objective of minimizing the total distance of data points to the centres of the clusters. In a cross-border regional analysis context, applying a k-means clustering approach would means dividing the grid cells included in the dataset into smaller clusters consisting of points that are close to each other. Lower k values mean that the cells in the dataset will be grouped in fewer clusters, resulting in larger cluster sizes. The higher the k value, the smaller the average size of each cluster becomes.

The clustering algorithm applied here takes only distance into account, so cluster formation depends only on the number of clusters to be formed (k) and the coordinates of each grid cell. Additional variables can be calculated for each cluster. For example, values of k between 10 and 150 were tested, resulting in average cluster sizes of 33000 cells to 2000 respectively. For each cluster at every k value selected, the average benefit in terms of cross-border potential accessibility was calculated and the ten clusters with the highest impact for every k value were mapped in Figure 52.

This process allows the comparison of the benefits across all clusters identified for all k values. Cluster size influences average impact, since the impact for each cluster component varies. Comparing a large number of potential combinations through k-means clustering is a relatively fast *brute force* algorithm that allows a ranking of different cluster size options using the same evaluation objective.

Figure 52. Top-10 clusters in terms of average improvement in potential accessibility, using various k values



11 Conclusions

Physical distance and –by extension- travel time on the road network can be one of the reasons that impede interaction across EU areas close to a land border. There are obviously several reasons why cross-border contacts are more limited than contacts at the same distance within the same country. Centuries of nation building across Europe have led to strong physical, cultural and linguistic borders that still persist even in a context of free movement inside the European Union. In many cases, border regions have been neglected at both national and international level. Priorities for transport infrastructure tended to favour more central zone or a few border crossings necessary for international trade. As a result, a large part of the internal EU land borders face the double burden of a difficult geographic situation and a lack of transport infrastructure investment to help reduce isolation.

Increasing cross-border collaboration is a policy objective at EU level that aims at improving the situation for border regions and exploiting the largely untapped potential of connecting neighbouring regions. A wide range of measures can be implemented in order to reduce the non-physical obstacles, such as linguistic, cultural or administrative barriers, but tangible, physical infrastructure may still play a role as a facilitator of collaboration. The methodology presented in this report concentrated on this aspect of connectivity, in particular on exploring ways to measure the quality of cross-border road networks and estimate the potential impact from targeted investment.

The population living in border regions is considerable, equal to 13% of the total EU population. In most cases there is a marked difference as regards the level of cross-border connectivity and accessibility that those zones enjoy in comparison with the respective levels on the national side. Even when taking the influence of local geographic barriers into account, cross-border links are significantly slower and longer than links within the same country.

Road accessibility of all border regions in EU28 (plus Norway and Switzerland) was assessed at grid level using three indicators addressing different policy perspectives. Network efficiency is compared to potential accessibility and the location indicator showing in practice the key advantages of network efficiency for the specific application.

Aggregating at the level of specific bilateral borders suggests that there is a distinction between older and newer EU member states, central and peripheral European countries or Western and Eastern European countries. At country level, the best performing road network in terms of both national and cross-border network efficiency can be found in the borders of the Netherlands, Belgium and Germany, while the worst performing road network can be found in the borders of the Balkan countries.

The results of the calculation of network efficiency at grid level are used to analyse specific cases. The combination of the results on network efficiency with the road network and population distribution indicates the reasons of low or high network efficiency and highlights the role of the road network. At a more detailed level, network efficiency is combined with data on population of border zone grid cells and destination settlements to identify areas of high interest, i.e. areas with relatively high population and low network performance. The highest concentration of such grid cells can be found in the borders of Czech Republic, but also in the borders between Romania and Bulgaria, Hungary and Slovakia, Hungary and Romania. Using further spatial information it is possible to explore the reasons of low network efficiency. A common reason of low network efficiency is the existence of natural barriers such as mountains (e.g. in the borders between Greece and Bulgaria, Czech Republic and Poland, Italy and France), rivers (e.g. in the borders between France and Germany, Romania and Bulgaria) or national parks (e.g. in the border between Slovakia and Poland). Furthermore, by including the information on both national and international network efficiency the areas in the border between Czech Republic and Poland in highest need for infrastructure investments are identified.

The methodology presented and the resulting dataset can provide useful input to policy makers and practitioners in terms of evaluating the current situation, prioritizing areas for intervention and assessing potential impacts. It should be noted, however, that the areas identified still need to be examined on a case by case basis in order to take into account all the relevant characteristics. The approach allows a focus on specific road links, settlements and physical characteristics, but other factors affecting cross-border collaboration should be also taken into account in additional steps of the analysis.

The overall approach can be extended to cover any area through adapting either the extension of the grid used (e.g. to analyse accessibility of wider zones or even the whole of the EU) or the size of the interest zone (to calculate indicators using different formulations or attraction factors). The results of this methodology can be easily combined with other policy support tools for tailored multi-criteria analysis.

References

Brakman, S., Garretsen, H., Van Marrewijk, C. and Oumer, A. 2012. The border population effects of EU integration. Journal of Regional Science, 52, 40-59.

Bruinsma, F. and Rietveld, P. 1998. The accessibility of European cities; theoretical framework and comparison of approached. Environment and Planning A, 30, 499–521.

Condeço-Melhorado, A. and Christidis, P. 2017. Road Accessibility in Border Regions: a Joint Approach. Networks and Spatial Economics, DOI: 10.1007/s11067-017-9362-1.

European Commission, Boosting growth and cohesion in EU border regions, Communication from the Commission to the Council and European Parliament, September 2017, Brussels, COM(2017) 534 final.

European Commission, Overcoming obstacles in border regions – Summary report on the online public consultation. Luxembourg, 2017.

Geurs, K.T. and Ritsema van Eck, J.R. 2001. Accessibility measures: review and applications. RIVM Report 408505 006, National Institute of Public Health and the Environment, Bilthoven.

Geurs, K.T. and van Wee, B. 2004. Accessibility of land-use and transport strategies: review and research directions. Journal of Transport Geography, 12, 127–140.

Gutiérrez, J. 2001. Location, economic potential and daily accessibility: An analysis of the accessibility impact of the high-speed line Madrid-Barcelona-French border. Journal of Transport Geography, 9, 4, 229–242.

Gutiérrez, J., González, R. and Gómez, G. 1996. The European high-speed train network: predicted effects on accessibility patterns. Journal of Transport Geography, 4,4, 227–238.

Gutiérrez, J. and Monzón, A. 1998. Accessibility, network efficiency, and transport infrastructure planning. Environment and Planning A, 30, 1337–1350.

Gutiérrez,J., Condeço-Melhorado,A., López, E. and Monzón, A. 2011. Evaluating the European added value of TEN-T projects: a methodological proposal based on spatial spillovers, accessibility and GIS. Journal of Transport Geography, 19, 840–850.

Jacobs-Crisioni, C. and Koomen, E. 2017. Population growth, accessibility spillovers and persistent borders: Historical growth in West-European municipalities. Journal of Transport Geography, 62, 80-91.

López, E., Gutiérrez, J. and Gómez, G. 2008. Measuring regional cohesion effects of large-scale transport infrastructure investments: an accessibility approach. European Studies, 16. 2, 277-301.

López, E., Monón, A., Ortega, E. and Mancebo Quintana, S. 2009. Assessment of Cross-Border Spillover Effects of National Transport Infrastructure Plans: An Accessibility Approach. Transport Reviews, 29, 4, 515-536.

Maggi, R., Masser, I. and Nijkamp, P. 1992. Missing networks in European transport and communications. Transport Reviews, 12, 4, 311–321.

Persyn, D. and Torfs, W. 2016. A gravity equation for commuting with an application to estimating regional border effects in Belgium. Journal of Economic Geography, 16, 155-175.

Rodrik, D. 2000. How far will international economic integration go? Journal of Economic Perspectives, 14, 177-186.

TomTom, 2015. MultiNet®, EUR 2015.12, Product Release Notes, www.tomtom.com/licensing. 2015 TomTom Global Content BV and TomTom North America, Inc.

List of figures

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications
You can download or order free and priced EU publications from EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

KJ-NA-29565-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2760/040279

ISBN 978-92-79-98380-1